2012

INVESTIGATION OF THE CRITICAL RESIDUES AT THE C-TERMINUS OF THE E.COLI PERIPLASMIC CHAPERONE, SURA

Brent Ferrell
University of Kentucky, bferrell3@gmail.com

STUDENT AGREEMENT:

I represent that my thesis or dissertation and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained and attached hereto needed written permission statements(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine).

I hereby grant to The University of Kentucky and its agents the non-exclusive license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless a preapproved embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

REVIEW, APPROVAL AND ACCEPTANCE

The document mentioned above has been reviewed and accepted by the student's advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student's dissertation including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

Brent Ferrell, Student

Dr. Yinan Wei, Major Professor

Dr. John Anthony, Director of Graduate Studies

INVESTIGATION OF THE CRITICAL RESIDUES AT THE C-TERMINUS OF THE *E.COLI* PERIPLASMIC CHAPERONE, SURA

THESIS

A thesis submitted in partial fulfillment of the
requirements for the degree of Master of Science in the
College of Arts and Sciences
at the University of Kentucky

By

Brent Ferrell

Lexington, Kentucky

Director: Dr. Yinan Wei, Assistant Professor of Chemistry

Lexington, Kentucky

2012

ABSTRACT OF THESIS

INVESTIGATION OF THE CRITICAL RESIDUES AT THE C-TERMINUS OF THE *E.COLI* PERIPLASMIC CHAPERONE, SURA

SurA is a molecular chaperone in the periplasm of *E.coli* and has been implicated in the maturation of outer membrane proteins (OMPs). SurA consists of four domains, but only two of them, namely the N and C-terminal domains, are necessary for chaperone function. Very little is known about which residues drive the interaction between the N and C-termini that facilitates normal activity. We mutated several conserved residues on the C-terminus and generated additive truncations to observe the effects of each on the fitness of the cell. We found one mutation E(408):A was sufficient to reduce SurA activity by 3-fold, but structural characterization of the mutated protein revealed little variation from wild-type SurA. Most notable, we found that when at least 10 residues are removed from the C-terminus, the protein is completely non-functional. We introduced a random peptide library to substitute these 10 residues and found that ~1.5% of all possible sequences in the library can restore SurA function to at least 50% activity. Moreover, we observed no pattern in the sequences of 26 different variants that were chosen and characterized. Here we show for the first time that SurA can tolerate many mutations at the C-terminus and still be active.

Brent Ferrell

April 26th, 2012

INVESTIGATION OF THE CRITICAL RESIDUES AT THE C-TERMINUS OF THE *E.COLI* PERIPLASMIC CHAPERONE, SURA

By

Brent Ferrell

_____Yinan Wei, Ph.D._____
Director of Thesis

_____John Anthony, Ph.D._____
Director of Graduate Studies

_____April 26th, 2012_____

Table of Contents

List of Tables

List of Figures

Chapter I: Introduction

1.1 Background and Significance

Over the last 40 years, antibiotic resistance among infectious bacteria has emerged and exploded into a worldwide problem that has resulted in once easily curable infections becoming untreatable. There are many factors that have contributed to the spread of antibiotic resistance. Among these are the lack of patient compliance with prescribed directives, rapid globalization, and the overuse and misuse of antibiotic agents [Mitscher, L. A., S. P. Pillai, et al. (1999)]. The influence of humans has undoubtedly accelerated the natural genetic variation among bacterial species and has selected for those, which are less susceptible to antibiotics even to the point of total resistance to every approved drug [Velayati, A. A., M. R. Masjedi, et al. (2009)]. As a result, the mechanisms of bacterial drug resistance have been an active field of study in the last few decades.

Bacteria have four main mechanisms of defense against antimicrobials: exclusion of the drug, chemical modification of the drug upon entry, variation of the drug target, and efflux of the drug. Furthermore, bacteria with intrinsic resistance to a chemical agent can pass this resistance on in two ways: horizontal and vertical gene transfer. Vertical gene transfer is accomplished by cell division and horizontal gene transfer can occur by a number of ways including: plasmid conjugation, transformation, or transduction [Mitscher, L. A., S. P. Pillai, et al. (1999)]. The high frequency of genetic mutations in bacteria and relative ease of gene transfer through a rapidly growing population has the potential to cause drastic changes in

the viability of that population over a short period of time. In order to combat the natural resistance that occurs from selection pressures imposed by the use of antibiotics, there must be a continual pursuit for novel approaches to keeping bacterial infection under control.

No novel class of antibiotics was discovered between 1962 and 2000. Since 2000, there have only been three major classes of antibiotics approved [Fischbach, M. A. and C. T. Walsh (2009)]. Many drugs that have been developed target cytoplasmic macromolecules, skirting around the natural mechanisms of resistance posed by bacteria. Consequently, many of them have been resisted by bacteria and rendered ineffective. A different approach might be to inactivate or competitively inhibit the mechanism of resistance, and then apply a conventional antibiotic to kill the cell. Much research has gone into understanding the ways in which bacteria acquire and retain resistance to antibiotics in hopes of developing ways to inactivate them. Of particular interest to this study is the exclusion mechanism, which prevents antibiotics from entering the cell at all.

1.2 Structure of Bacterial Cell Membrane

Most bacteria are divided into two categories based on their ability to retain or exclude the Gram stain. Those which retain Gram stain are known as Gram-positive bacteria and those which exclude it are Gram-negative bacteria. There exist distinct differences in the cell membrane structure of these two types of cells. Both cell types have a symmetric inner membrane composed mainly of phospholipids

[Baron, S. (1996)]. Outside of that, Gram-positive cells have a thick peptidoglycan layer, with teichoic and lipoteichoic acids dispersed throughout the matrix.

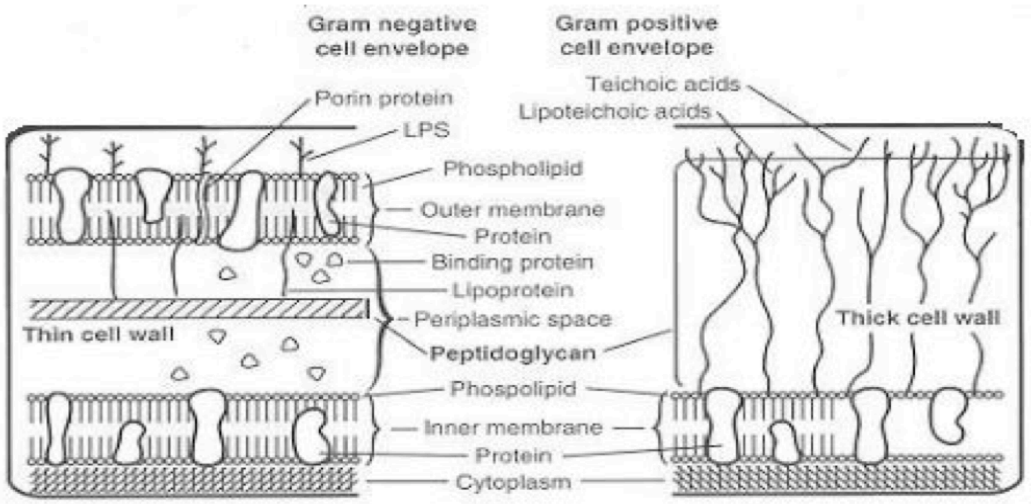

Figure 1.1 Gram-negative cell envelope (left) and Gram-positive cell envelope (right) [taken from Baron, S. (1996)]

Gram-negative cells have only a thin peptidoglycan layer, but are guarded by an asymmetric outer membrane, composed of a complex mixture of phospholipids and lipopolysaccharide (LPS). It is the extra outer membrane in gram-negative cells that poses a significant barrier to the entry of small molecules. Consequently, when this barrier is compromised, the cell becomes more susceptible to antibiotics. One factor that can affect the outer membrane integrity is the protein content. Low outer membrane protein (OMP) content leads to a leaky membrane, allowing passage of small molecules normally excluded [Rouviere, P. E. and C. A. Gross

(1996)]. Many factors can contribute to the level of outer membrane protein content, thereby affecting the permeability of the cell to antibiotics. Periplasmic chaperones are implicated in the biogenesis and transport of outer membrane proteins [Rouviere, P. E. and C. A. Gross (1996), Lazar, S. W. and R. Kolter (1996), Rizzitello, A. E., J. R. Harper, et al. (2001), Bos, M. P., V. Robert, et al. (2007)]. The focus of this study is how periplasmic molecular chaperone activity affects the susceptibility of *E .coli* cells to antibiotic agents.

1.3 Periplasmic Molecular Chaperones of *E. coli*

The periplasm is the space in Gram-negative bacterial cells that separates the inner membrane from the outer membrane. The periplasm is an oxidizing, energy-deficient environment that is more crowded with proteins and polysaccharides than the cytoplasm [Pluckthun, A, Wulfing, C. (1994)]. Furthermore, the periplasm is an extremely dynamic environment, constantly exchanging ions and small molecules, and responding to extracellular stress such as toxic molecules, pH changes, and heat. Proteins destined for the outer membrane must pass through the periplasm, often with the aid of chaperones. In prokaryotic cells, membrane proteins are synthesized by ribosomes associated with the plasma membrane. Similar to secretory and inner membrane (IM) proteins, the translocation of the nascent polypeptide chain of OMP across the IM is assisted by translocons, which are protein complexes creating a water-filled channel in the IM [White, S. H. et. al. (2009), Gold, V. A. M., Duong, F. and Collinson, I. (2007), Lizak, B., Csala, M., et. al. (2008), Rusch, S. L. and Kendall, D. A.

(2007)]. The process of OMP translocation through the periplasm and insertion into the OM is poorly defined. However, it is becoming clear that the process requires the assistance of periplasmic chaperones, proteins that bind to and stabilize the nascent OMP to prevent misfolding and precipitation, and potentially also assist the insertion of the proteins to their final destination in the OM [Pluckthun, A, Wulfing, C. (1994), Bos, M. P., V. Robert, et al. (2007), Ruiz N., Kahne D., Silhavy T.J. (2006), Missiakas, D., J. M. Betton, et al. (1996), Duguay A. R., Silhavy T. J. (2004), Bernstein, H.D. (2000), Sijbrandi, R., Urbanus, M.L., et al. (2003)].

Several periplasmic molecular chaperones have been identified, including DegP, Skp, FkpA and SurA [Mogensen J. E., and Otzen D. E. (2005), Sklar J. G., Wu T., et. al., (2007), Stymest K. H., and Klappa P. (2008)]. Unlike the cytoplasmic chaperones, the periplasmic chaperones operate without the consumption of energy, as ATP is not present in the periplasm. In addition to molecular chaperones, several peptidyl-prolyl isomerases and at least six disulfide bond reductases and isomerases have been identified in the periplasm [Behrens, S., R. Maier, et al. (2001)]. Despite the important role of chaperones, only a few have been discovered and examined. Among the chaperones, the interactions of Skp and several OMP substrates have been studied most extensively [Qu, J., C. Mayer, et al. (2007), Chen R. and Henning U. (1996), Harms N., Koningstein G., et. al. (2001), Schafer U., Beck K. and Muller M. (1999), Bulieris P. V., Behrens S., et. al. (2003)].

Cells lacking Skp are viable and show no signs of sickness [Rizzitello, A. E., J. R. Harper, et al. (2001)] but express lower levels of OMPs [Chen R. and Henning U.

(1996)]. In addition, a study has indicated that Skp interacts with the outer membrane porin, PhoE, at the inner membrane interface during translocation from the cytoplasm to the periplasm [Harms N., Koningstein G., et. al. (2001)]. As a result of this study, it is believed that Skp is the first molecular chaperone to interact with non-native OMPs in the periplasm. The mechanism of this interaction is purported to be the result of hydrophobic and electrostatic interactions between a homotrimer of Skp and the unfolded OMP [Qu, J., C. Mayer, et al. (2007)]. However, the exact binding site for Skp remains elusive.

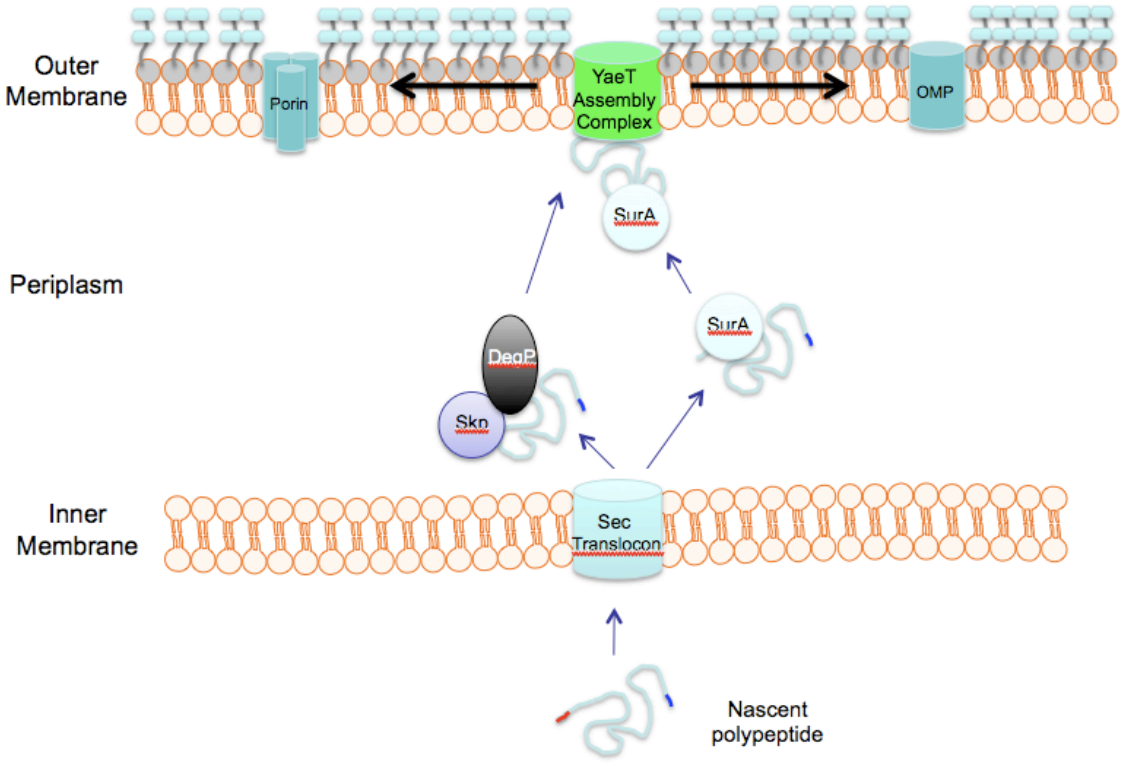

Figure 1.2 OMP biogenesis involving periplasmic molecular chaperones SurA, Skp, and DegP. Nascent OMP polypeptides emerge from ribosomes and are translocated out of the cytoplasm by the Sec system via an C-terminal signal peptide (red). In the periplasm, polypeptides are either met first by Skp/DegP or SurA. OMPs that are encoded with a N-terminal signal peptide are assembled into the outer membrane by the YaeT complex. It is known that SurA interacts with the YaeT complex during the process of OMP assembly [Sklar, J. G., T. Wu, et al. (2007)]. By deduction, it is assumed that Skp/DegP also must interact with this complex when SurA is not present because certain OMPs must be assembled by YaeT.

SurA is a periplasmic protein in *E. coli* that has dual functions as a molecular chaperone and peptidyl prolyl *cis/trans* isomerase (PPIase) [Rouviere, P. E.; Gross, C. A. (1996)]. It has been purported to be the primary chaperone among the three major periplasmic chaperones: DegP, Skp, and SurA [Sklar, J. G., T. Wu, et al. (2007)]. SurA was originally described as a protein necessary for stationary phase growth [Tormo, A., M. Almiron, et al. (1990)], but is not vital for growth under normal conditions [Justice, S. S., D. A. Hunstad, et al. (2005)]. At low temperatures, however, cells lacking SurA exhibit growth defects [Denoncin, K., D. Vertommen, et al. (2010)]. In addition, studies have shown that *surA*-null mutants are more susceptible to certain antibiotics, bile salts, and SDS than *surA*-containing strains, indicating a defective outer membrane [Denoncin, K., D. Vertommen, et al. (2010), Watts, K. M. and D. A. Hunstad (2008), Justice, S. S., D. A. Hunstad, et al. (2005), Behrens, S., R. Maier, et al. (2001), Rouviere, P. E. and C. A. Gross (1996), Lazar, S. W. and R. Kolter (1996)]. Furthermore, *surA*-null strains form slightly mucoid colonies and are easily lysed during stationary phase, consistent with the phenotype with a defective outer membrane [Lazar, S. W. and R. Kolter (1996)]. Similar to Skp, surA-deficient cells show significantly lower levels of OMPs when compared with wild-type (WT) cells [Rouviere, P. E. and C. A. Gross (1996)]. Indeed, lower levels of OMPs could lead to a leaky outer membrane allowing the passage of small molecules normally excluded.

Both SurA and Skp were demonstrated to have functional redundancy in *E.* coli [Rizzitello, A. E., J. R. Harper, et al. (2001)]. A single deletion of *surA* results in an

up-regulation of both DegP(~30-fold) and Skp(~3-fold), by Western blot analysis[Palomino, C., E. Marin, et al. (2011)]. A double deletion of *surA* and *skp* genes produces a synthetically lethal phenotype in rich media, resulting in bacteriostatic cells. In addition, these cells exhibit lowered levels of OMPs and changes to the cell morphology when compared to WT cells. Absence of DegP, another periplasmic chaperone/protease, along with SurA causes a bactericidal phenotype in *E. coli.* This, in contrast to the bacteriostatic effect seen in the double *surA skp* mutant, is due to the lack of DegP protease activity in the periplasm. However, simultaneous deletion of *skp* and *degP* does not result in a lethal phenotype [Rizzitello, A. E., J. R. Harper, et al. (2001)]. Based on these findings, it is possible that there are two separate chaperone pathways, which are functionally redundant: one pathway involving SurA, and the other involving both DegP and Skp.

1.4 Characterization of SurA

Following the earliest studies on SurA, showing that it was required for stationary phase growth under special conditions, many pleiotropic defects were found in the cell envelope of *surA*-null mutants. When sequenced, the *surA* gene was found to have regions of high sequence homology with the cytoplasmic peptidyl-prolyl isomerase (PPIase), parvulin. These findings, coupled with the periplasmic location of SurA, indicated that SurA could assist the folding of extracytoplasmic proteins. To confirm that SurA assisted outer membrane protein folding, several OMPs were tested for trypsin sensitivity in WT and *surA*-null mutants. OMPs

extracted from the *surA*-null strain were more sensitive to trypsin digestion, indicating mal-folded OMPs [Lazar, S. W. and R. Kolter (1996)]. This was the first evidence showing that SurA is involved in outer membrane protein folding. A subsequent study revealed that SurA has PPIase activity, as evidenced by its affinity for artificial proline-containing substrates and catalysis of proline isomerization. This study also pointed out that SurA is involved in OMP maturation, and OMP levels are lowered due to reduced genetic transcription (σ^E induction) and improper folding in *surA*-null mutants [Rouviere, P. E. and C. A. Gross (1996)]. The early studies on SurA were grounded in observing phenotypic changes between WT and *surA*-null mutants. Later studies have been focused on the details of SurA function and which regions of the protein are implicated in the function.

SurA is composed of 428 amino acid residues with four domains: N-terminal domain (N) (21-175), PPIase 1(P1) domain (176-278), PPIase II(P2) domain (286-384), and C-terminal domain(Ct) (385-428) [Bitto, E. and D. B. McKay (2002)]. It has been shown that the chaperone activity of SurA operates independent of its PPIase activity. If either or both of 'P' domains in SurA are deleted, the protein can still function *in vivo* with almost complete activity, as evidenced by plating on SDS-EDTA, novobiocin, and measurements of σ^E induction [Behrens, S., R. Maier, et al. (2001), Watts, K. M. and D. A. Hunstad (2008)]. However, no individual domain is sufficient to rescue the *surA*-null phenotype and both the N and C-terminal domains are indispensible for proper function. To assess whether the N and C-terminal domains alone exhibited chaperone activity *in vitro*, a N-CtΔ(P1-P2) mutant was tested for its ability to prevent aggregation of unfolded citrate synthase. Indeed, this

mutant was capable of inhibiting thermal aggregation of citrate synthase, a classic assay for molecular chaperones [Behrens, S., R. Maier, et al. (2001)]. This result pointed out that the N and C-terminal domains are largely responsible for the chaperone activity of SurA.

The crystal structure of SurA has been solved, revealing four distinct domains, consistent with sequence alignment result with other peptidyl-prolyl isomerases [Bitto, E. and D. B. McKay (2002)]. The N and C-terminal domains, along with the P1 domain form the "core module" of the protein, with the satellite P2 domain separated by two polypeptide linkers of ~25-30 angstroms. In the core module, 60% of the C-terminal domain surface area interface is in contact with the N-domain, showing high association between the two interfaces. In addition, topology images of the core domain show a deep crevice formed mostly by helices of the N-terminal domain. This crevice was observed in the crystal structure of SurA, where a polypeptide (residues 153-164) binds in the neighboring core module region for approximately 15 angstroms in length. The short peptide only fills ~30% of the channel, which suggests that much longer polypeptides could be accommodated in this binding pocket [Bitto, E. and D. B. McKay (2002)].

SurA is substrate-specific and was found to bind to an *in-vitro* synthesized unfolded porin 50-fold more efficiently than other proteins of similar size [Behrens, S., R. Maier, et al. (2001)]. By using phage display, it was found that SurA preferentially binds proteins with an Ar-X-Ar motif, where Ar is an aromatic residue and X is any other residue [Bitto, E. and D. B. McKay (2003)]. This motif is

characteristic of integral OMPs in *E. coli,* suggesting that the primary targets of SurA are unfolded OMPs. Several heptapeptides, selected by phage display, were tested for binding with both SurA and SurA(ΔP2) and were found to bind with low micromolar affinity. The SurA(ΔP2) fusion had even higher (less than 2-fold) affinity for the short peptides than full-length SurA, further demonstrating that the P2 domain is not involved in chaperone activity. Futher, it was found that one of the heptapeptides (FTYMPPV) competes with unfolded OmpG and OmpF for the binding with SurA, indicating that the selected heptapeptides share overlapped binding site on SurA with unfolded OMPs.

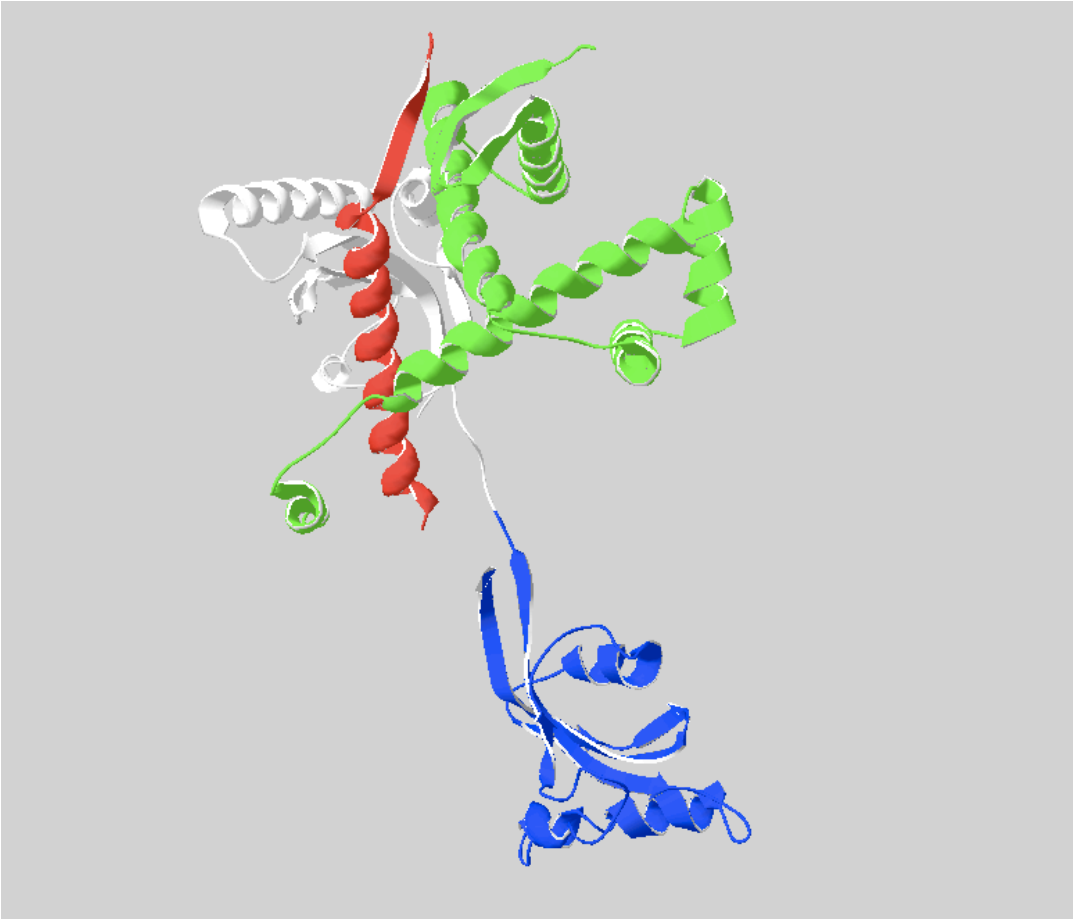

Figure 1.3 Crystallographic structure of SurA. The N-terminal domain is shown in green, the P1 domain is white, the P2 domain is blue, and the C-terminal domain is red. Structure was generated from the Swiss Protein Data Bank, ID: 1M5Y [Bitto, E. and D. B. McKay (2002)].

A more detailed study on SurA binding focused on the interaction of SurA with specific regions of three OMPs (OmpF, LamB, and OmpA). Short (13-mer) peptides were synthesized of the entire protein sequence for each OMP and immobilized onto a cellulose plate. SurA was found to bind the regions of the OMPs with the most aromatic content. 73% of the good (>60%) binders and 62% of the weak (>40%, <60%) binders contained the Ar-X-Ar or Ar-Ar motif. Of the good and weak binders, the majority of them show a preference for generating beta-strand secondary structure, based on predictions by the computer program, PsiPred [Xu, X. H., S. Y. Wang, et al. (2007)]. In addition, SurA lacking the P2 domain bound to the same set of peptides as full-length SurA, further confirming that the core domain of SurA is the active chaperone region [Hennecke, G., J. Nolte, et al. (2005)].

The exact binding site on SurA has not been accurately determined, but it has been suggested to exist in the P1 domain. This assertion is based on two findings: 1) Crystal structures show that both the C-terminal phage displayed dodecapeptide, NFTLKFWDIFRK [Xu, X. H., S. Y. Wang, et al. (2007)] and the septapeptide, WEYIPNV, [Bitto, E. and D. B. McKay (2003)] bind to the P1 domain of SurA. 2) The binding of the dodecapeptide and septapeptide with the isolated P1 domain of SurA was found to be five and 50-fold higher, respectively, than with full-length SurA [Xu,

X. H., S. Y. Wang, et al. (2007)]. The sequence of the dodecapeptide defies the previously determined consensus motif Ar-X-Ar, but still contains 33% aromatic content, in agreement with the notion that SurA recognizes polypeptide sequences rich in aromatic residues. These crystal structures, and one other from the same study (SurA(ΔP2) complexed with the C-peptide), exist as the only conclusive evidence we have of a binding site for SurA. However, if the P1 domain were the binding site of unfolded OMPs, then the *surA*-null mutants containing *surA* constructs lacking the P1 domain would not exhibit a WT phenotype. It should be noted that the *surA* constructs N-P1-Ct and N-P2-Ct are slightly more active *in vivo* and *in vitro* than the N-Ct alone [Behrens, S., R. Maier, et al. (2001)]. Thus, the P1 and P2 domains might play an assistant role in the chaperone activity of SurA. Since the P1 domain alone is insufficient to provide chaperone activity, but recognizes peptide domains characteristic of OMPs, it could be involved in substrate recognition. It was proposed that the crevice formed by the N-domain provides a region for large segments of OMPs to bind with broad specificity, while the P1 domain enhances selectivity of the SurA-OMP complex [Xu, X. H., S. Y. Wang, et al. (2007)]. This would explain why the N-Ct domains are able to function as a chaperone alone, but not as well as the N-P1-Ct domains together.

The specificity of SurA toward certain OMPs was confirmed in a proteomics study, where 64 OMPs, including 23 beta-barrel proteins, were assessed for relative abundance in the outer membrane (OM) fraction of WT and *surA*-null mutants. Of the 23 beta-barrel proteins, eight of them were found to be negatively affected by lack of SurA. Consistent with previous results, the most studied SurA substrates,

OmpF, OmpA, and LamB, were among this group. These three OMPs constitute the majority of OMPs in the outer membrane [Sklar, J. G., T. Wu, et al. (2007)]. However, the levels of only two proteins, LptD and FhuA, were reduced without a loss in genetic transcription. This suggests that LptD and FhuA could be 'true' substrates of SurA, with normal levels of mRNA, but absence from the OM fraction in *surA*-null mutants [Vertommen, D., N. Ruiz, et al. (2009)]. Other OMPs also bind to SurA [Hennecke, G., J. Nolte, et al. (2005)], but it is difficult to delineate whether the reduction of these OMPs in the OM fraction is a result of the lack of SurA or reduced mRNA levels. It is most probable that the reduction of mRNA is an indirect result (via σ^E induction) of the accumulation of unfolded and aggregated OMPs due to the lack of SurA in the periplasm.

Recently, SurA was implicated as a chaperone for the fimbrial usher, FimD [Justice, S. S., D. A. Hunstad, et al. (2005)]. FimD is involved in the polymerization of multimeric, adhesive pili, an essential component for the pathogenesis of *E. coli* by mediating attachment to endothelial tissue. Cells lacking SurA have much lower levels of FimD, as shown by Western blot with anti-FimD antibody. The downstream effects of less FimD were also observed, where *surA*-null uropathogenic *E.coli* (UPEC) cells were less able to mediate invasion of human endothelial cells *in vitro*. Both decrements were restored to 50% or higher of wild-type levels with plasmid constructs of full-length *surA*, *surA*N+Ct, and *surA*N-(P1 or P2)-Ct [Watts, K. M. and D. A. Hunstad (2008)]. A subsequent study confirmed these findings in more detail by determining that FimD assembly in the outer membrane can be specifically attributed to the presence of SurA. An important effect of the

absence of SurA in the cell is the upregulation of DegP, a periplasmic protease. One could conclude that reduced FimD in the outer membrane is just a result of increased levels of DegP degrading unfolded FimD. However, in the absence of DegP and SurA, the levels of folded FimD still dropped sharply, while the levels of unfolded FimD accumulated. Thus, DegP is responsible for degrading unfolded FimD, but SurA is required for proper insertion of folded FimD. It should be noted that strains lacking only DegP exhibit normal levels of FimD in the outer membrane [Palomino, C., E. Marin, et al. 2011].

1.5 Summary

Many studies have indicated the importance of SurA as a chaperone in the periplasm. SurA is an important component in OMP biogenesis and cells lacking SurA are much more susceptible to membrane perturbants than wild-type strains. Despite all that is known about SurA, much is left a mystery. The exact binding site on SurA for full-length OMPs is only postulated based on molecular modeling. Furthermore, there have been no in-depth mutational studies to investigate the critical residues of SurA. My primary aim in this work was to explore the importance of the C-terminus of SurA. I accomplished this in two ways: 1) 15 single/double/triple site-directed mutations were generated and these mutants were tested for activity through a minimum inhibitory concentration (MIC) assay of novobiocin, 2) 26 mutants were generated where the last 10 amino acids of SurA were replaced by random 10-mer peptides and tested for MIC with novobiocin and

SDS-EDTA. A secondary aim was to mutate all of the aromatic residues at the N and C-termini to alanine in single, double and triple mutation combinations, and test the MIC of novobiocin. Of the first 15 mutants generated from site-directed mutations, only one mutant (E408:A) showed a significant (>3-fold) reduction in MIC from WT. The resulting SurA protein ($SurA_{E408A}$) was purified and characterized using circular dichroism, fluorescence spectroscopy, and fluorescence polarization (binding with OmpF). Of the aromatics that were mutated, only two of the triple mutants, Y(120, 398, 422):A and Y(125, 398, 422):A displayed low activity. The study of the SurA C-terminal tail uncovered perhaps the most interesting and informative results. Of the 26 mutants selected, the MIC range for novobiocin was 500-1200 and 1.5%->5.0% SDS (with 1mM EDTA), indicating a wide range of SurA activity. However, the sequences of these mutants revealed very little homology to wild-type or even among the mutants of similar activity. Furthermore, the expression levels of SurA have little to do with the activity, as indicated by the MIC assays. Three of the mutants were chosen (two of high activity and one of modest activity) for protein characterization. The results indicated that the protein tertiary structure differed greatly from wild-type. Taken together, it can be concluded that SurA is very tolerant to mutations of its sequence, which is consistent with its role in the interaction with nascent OMPs of various sequences and structures.

Chapter II: Materials and Methods

2.1 Materials

Sodium hydroxide, glycine, sucrose potassium hydroxide, Tween-20, ethylenediaminetetraacetic acid (EDTA), calcium dichloride dihydrate, kanamycin sulfate, tris-hydroxymethyl aminomethane (TRIS), sodium dodecyl sulfate (SDS), sodium chloride, agar (for plates), sodium phosphate (dibasic), tetramethylethyelenediamine (TEMED), glycerol, acrylamide, bovine serum albumin (BSA), and ammonium persulfate were purchased from BioWorld Inc. (Atlanta, GA). Luria-Bertani Broth, tryptone, yeast extract, and ampicillin were purchased from Boston Bioproducts (Boston, MA). Novobiocin, Sodium N-laurylsarcosine and Fluorescein-isothiocyanate (FITC) were purchased from Sigma-Aldrich (St. Louis, MO). Ni-NTA agarose beads were purchased from Qiagen (Valencia, CA). Gel extraction and plasmid mini-prep kits, Semi-dry blotter, PCR tubes, nitro-blue tetrazolium (NBT), 5-bromo-4-chloro-3'-indolyphosphate (BCIP),and agarose HS (for gels) were purchased from Denville (Metuchen, NJ). Ethanol was purchased from Amresco (Solon, OH). Methanol was purchased from VWR (Radnor, PA). Water was purified using a Milli-Q® water purification system by Millipore. PVDF transfer membranes and Amicon centrifugal filter units were purchased from Millipore (Billerica, MA).

2.2 Cloning, Expression, and Purification

E.coli surA-null strain JW0052 was obtained from the Coli Genetic Stock Center at Yale University, where *surA* was deleted from the genomic DNA and replaced with a kanamycin-resistance cassette. *surA* gene was amplified using PCR from *E.coli* genomic DNA and cloned into pMalIII plasmid using Nde1 and EcoR1 restriction sites [Meng Zhong, unpublished result].

Plasmid-encoded SurA was transformed in *E. coli* ER2566 strain for expression. Protein expression and purification was conducted as described [Bitto, E. and D. B. McKay (2002)]. Centrifugation was done using the Allegra 25R Centrifuge (Beckman Coulter Indianapolis, IN).

2.3 Site-Directed Mutagenesis and Primer Sequences

Single residues in *surA* were altered by PCR, using the Quik change™ protocol. Primer pairs are listed in Table 1. WT *surA* gene, encoded on plasmid pMalIII, was used as the template, except in double mutants where one of the two codons was already mutated in the template. PCR parameters were as follows: 60 s at 95°C, followed by 22 cycles of 45 s at 95°C, 60 s, at 58°C and 13 min. at 68°C. , ending with 10 min. at 68°C. PCR was performed on a MJ mini Personal Thermal Cycler (BioRad, Hercules, CA). Following PCR, Dpn1 restriction enzyme was added to the PCR reaction mixture to digest any methylated parent strands. Next, the PCR product was transformed into chemically competent XL Blue cells. Transformation was accomplished by adding ~0.1µg of plasmid DNA to competent cells on ice and incubating for 30 min. Next, cells were heat-shocked at 42°C for 60 s, followed by

immediate incubation on ice for 5 min. 1 mL of LB broth was then added to the cells and incubated at 37°C for at least 45 min. Cells were collected by centrifugation, re-distributed in ~100uL of LB, and spread onto ampicillin-containing agar plates. Plates were incubated overnight at 37°C. At least 4 colonies from the plate were picked, shaken at 250 rpm at 37°C to late-log phase in LB before harvesting the cells and purifying the plasmid. Plasmid was purified using the plasmid mini-prep kit from Denville Scientific. Following purification, the plasmid was sent for sequencing (Operon, Huntsville, AL). All sequences analyzed using CLC Sequence Viewer 6 (CLC bio, Aarhus, Denmark). Centrifugation was done using the 260D microcentrifuge (Denville, Metuchen, NJ).

Table 2.1 Primers used in this study

W(413):A	CGGAAGAAGCAGCAAGCGCCATGCAGGAACAACGTGCC
	GGCACGTTGTTCCTGCATCGGGCTTGCTGCTTCTTCCG
R(418):G	GCTGGATGCAGGAACAAGGTGCCAGCGCCTACGTTAAA
	TTTAACGTAGGCGCTGGCACCTTGTTCCTGCATCCAGC
A(421):W	GCAGGAACAACGTGCCAGCTGGTACGTTAAAATCCTGAGC
	GCTCAGGATTTTAACGTACCAGCTGGCACGTTGTTCCTGC
I(425):S	GCCAGCGCCTACGTTAAATCCCTGAGCAACGGTGGTTC
	GAACCACCGTTGCTCAGGGATTTAACGTAGGCGCTGGC
E(416):A	GCAGCAAGCTGGATGCAGGCACAACGTGCCAGCGCC
	GGCGCTGGCACGTTGTGCCTGCATCCAGCTTGCTGC
K(405):A	CGCATGCTGATGAACCGTGCGTTCTCGGAAGAAGCAGCA
	TGCTGCTTCTTCCGAGAACGCACGGTTCATCAGCATGCG
N(403):A	GCATACCGCATGCTGATGGCCCGTAAGTTCTCGGAAGAAG
	CTTCTTCCGAGAACTTACGGGCCATCAGCATGCGGTATGC
E(408):A	GATGAACCGTAAGTTCTCGGCAGAAGCAGCAAGCTGGATG
	CATCCAGCTTGCTGCTTCTGCCGAGAACTTACGGTTCATC
Q(415):A	GAAGCAGCAAGCTGGATGGCGGAACAACGTGCCAGCGCC
	GGCGCTGGCACGTTGTTCCGCCATCCAGCTTGCTGCTTC
F(406):A	ATGCTGATGAACCGTAAGGCCTCGGAAGAAGCAGCAAGC
	GCTTGCTGCTTCTTCCGAGGCCTTACGGTTCATCAGCAT
E(409):A	GAACCGTAAGTTCTCGGAAGCAGCAGCAAGCTGGATGCAG
	CTGCATCCAGCTTGCTGCTGCTTCCGAGAACTTACGGTTC
S(407):A	GCTGATGAACCGTAAGTTCGCGGAAGAAGCAGCAAGCTGG
	CCAGCTTGCTGCTTCTTCCGCGAACTTACGGTTCATCAGC
Y(398):A	GCGCAGAAAGATCGTGCAGCGCGCATGCTGATGAACCG
	CGGTTCATCAGCATGCGCGCTGCACGATCTTTCTGCGC
Y(422):A	CAACGTGCCAGCGCCGCCGTTAAAATCCTGAGCAACGG
	CCGTTGCTCAGGATTTTAACGGCGGCGCTGGCACGTTG

Table 2.1 (continued)

Y(128):A	GGACTGAACTACAACACCGCTCGTAACCAGATCCGC
	GCGGATCTGGTTACGAGCGGTGTTGTAGTTCAGTCC
Y(120):A	GCGCAGCCGTCTGGCTGCCGATGGACTGAACTACAAC
	GTTGTAGTTCAGTCCATCGGCAGCCAGACGGCTGCGC
Y(422):A	CAACGTGCCAGCGCCGCCGTTAAAATCCTGAGCAACGG
	CCGTTGCTCAGGATTTTAACGGCGGCGCTGGCACGTTG
Y(125):A	GCTTACGATGGACTGAACGCCAACACCTATCGTAACCAG
	CTGGTTACGATAGGTGTTGGCGTTCAGTCCATCGTAAGC
W(413):G R(418):G	GGGATGCAGGAACAAGGTGCCAGCGCCTACG
	CGTAGGCGCTGGCACCTTGTTCCTGCATCCC
S(407):A E(408):A	GCTGATGAACCGTAAGTTCGCGGCAGAAGCAGCAAGCTGG
	CCAGCTTGCTGCTTCTGCCGCGAACTTACGGTTCATCAGC
E(408):A E(409):A	GAACCGTAAGTTCTCGGCAGCAGCAGCAAGCTGGATGCAG
	CTGCATCCAGCTTGCTGCTGCTGCCGAGAACTTACGGTTC
Y(120):A Y(125):A	GCTGCCGATGGACTGAACGCCAACACCTATCGTAACCAG
	CTGGTTACGATAGGTGTTGGCGTTCAGTCCATCGGCAGC
Y(120):A Y(128):A	GGACTGAACTACAACACCGCTCGTAACCAGATCCGC
	GCGGATCTGGTTACGAGCGGTGTTGTAGTTCAGTCC
Y(125):A Y(128):A	GGACTGAACGCCAACACCGCTCGTAACCAGATCCGCAAAG
	CTTTGCGGATCTGGTTACGAGCGGTGTTGGCGTTCAGTCC
Y(398):A Y(422):A	CAACGTGCCAGCGCCGCCGTTAAAATCCTGAGCAACGG
	CCGTTGCTCAGGATTTTAACGGCGGCGCTGGCACGTTG
4AA C-terminal deletion	CAGCGCCTACGTTAAATAACTGAGCAACGGTGGTTCTC
	GAGAACCACCGTTGCTCAGTTATTTAACGTAGGCGCTG
5AA C-terminal deletion	GTGCCAGCGCCTACGTTTAAATCCTGAGCAACGGTGGT
	ACCACCGTTGCTCAGGATTTAAACGTAGGCGCTGGCAC
10AA C-terminal deletion	AGCTGGATGCAGGAACAGCGCTAAAGCGCCTACGTTAAAATCCTG
	CAGGATTTTAACGTAGGCGCTTTAGCGCTGTTCCTGCATCCAGCT

Table 2.1 (continued)

11AA C-terminal deletion	GCTGGATGCAGGAACAATAAGCCAGCGCCTACGTTAAA
	TTTAACGTAGGCGCTGGCTTATTGTTCCTGCATCCAGC
20AA C-terminal deletion	GCTGATGAACCGTAAGTTCTAAGAAGAAGCAGCAAGCTGG
	CCAGCTTGCTGCTTCTTCTTAGAACTTACGGTTCATCAGC

2.4 Circular Dichroism (CD) Spectroscopy

CD was performed on a JASCO J-810 spectrometer (JASCO, United Kingdom) with 1 nm bandwidth, a 0.1 cm pathlength cuvette for secondary CD, and a 1.0cm pathlength for tertiary CD. The sample was dialyzed three times overnight into a low salt buffer (25 mM Phosphate buffer, 100mM NaCl pH 7.5) before the CD measurement. Blank scans were performed with the exterior dialysis buffer and subtracted from the measured data.

2.5 Fluorescence spectroscopy

Fluorescence polarization experiment and wavelength scans were performed with a Perkin Elmer LS-55 fluorescence spectrometer (Perkin Elmer, Waltham, Massachusetts). Scans were performed at room temperature. For the Trp fluorescence measurement, the protein was dialyzed overnight into a Phosphate buffer (20 mM, pH 8.0). The excitation wavelength was 280 nm.

For the fluorescence polarization studies of fluorescein isothiocyanate (FITC) labeled SurA with outer membrane proteins, the excitation and emission wavelengths were 492 and 515 nm, respectively. SurA was first labeled with fluorescein isothiocyanate (FITC), the molar ratio of SurA and FITC is 1:2. Unreacted

FITC was quenched using Tris and removed through dialysis. FITC-SurA concentration was determined by both the Bradford assay and fluorescence. Unfolded OMP was purified in the presence of 8 M urea. A solution containing 1.3 μM FITC-SurA was loaded in the cuvette, while small aliquots of unfolded OmpF were titrated into the FITC-SurA solution. A one-to-one binding model fit the data reasonably well.

2.6 SurA Activity Assay

The minimum inhibitory concentration (MIC) of primarily novobiocin was used to assess the activity of SurA in mutants. The MIC is the lowest amount of antibiotic needed to completely inhibit growth of the bacteria. Δ*surA* JW0052 *E. coli* cells were transformed with *surA*-containg pMalIII plasmid containing ampicillin resistance gene. Following transformation, individual colonies were picked and grown to an optical density (OD) of 0.1-0.15. Cells were diluted 10-fold and a 1uL spot was applied to the agar plates, followed by overnight incubation at 37°C. Agar plates were prepared with 1.5% agar and 1X LB along with various concentrations of antibiotics by adding the appropriate amount of antibiotic stock to 20mL of liquid agar mixture and allowing them to cool prior to inoculation. Following overnight incubation, the plates would be observed for colony formation. The concentration of drug at which no colonies formed was deemed the MIC for that mutant.

2.7 Random Peptide Insertion at C-terminus of SurA

A stop codon was introduced into the WT *surA* gene at position 419(Quik change™ protocol), resulting in a 10 amino acid deletion at the C-terminus during translation (denoted EC-1).

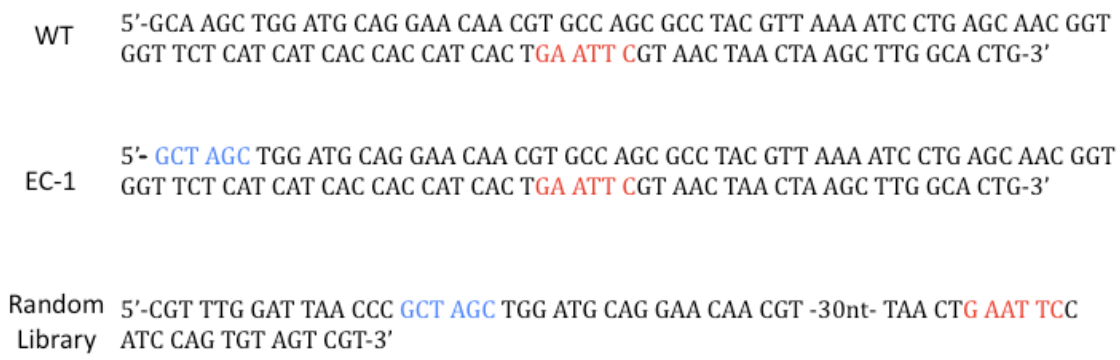

Figure 2.1 Gene sequences of WT *surA* and EC-1 *surA* construct beginning with Nhe1 cutting site (blue) and containing intrinsic EcoR1 cutting site (red).

The last 48 nucleotides of the EC-1 *surA* gene were removed by double digestion using a designed NheI cutting site (Quik change™ protocol) and an intrinsic EcoR1 cutting site in the pMalIII vector. An HPLC-purified DNA library (5'-CGTTTGGATTAACCCGCTAGCTGGATGCAGGAACAACGT-30nt-TAACTGAATTCCATCCAGTGTAGTCGT-3') containing a random 30-mer oligonucleotide sequence (IDT, Coralville, IO) flanked by primers corresponding to the C-terminal region of *surA* was amplified by PCR. The PCR product was column-purifed and ligated into the EC-1 vector at 16°C for at least 6 hours. The ligation product was transformed into the *surA*-null strain, JW0052. The transformed cells were plated on solid LB-agar media containing 150µg/mL novobiocin and allowed

to grow for 48 hours at 37°C. The resulting colonies were grown in LB media and the *surA*-containing plasmid DNA was purified using a plasmid-MiniPrep kit (Denville, Metuchen, NJ). Following sequencing, the plasmid DNA was re-transformed into the *surA*-null strain and at least 3 colonies from each mutant were picked for analysis of MIC using novobiocin.

2.8 6x-Histidine Tag Insertion For Purification of Mutants

Three mutants, M2-1, M2-7, and M2-9 were chosen for protein purification and expression. Two rounds of PCR were used to sequentially add in a 6X-His-tag. The F-primer for both rounds containing an intrinsic BamH1 cutting site was AAAGAGTTCTCTCAGGATCCAGGCTC and the first reverse primer varied with each mutant sequence: M2-1: GTGGTGATGATGAGAACCACCCATGCTCGGAGTAGCAACC, M2-7: GTGGTGATGATG AGAACCACC GGCGTTATGTCTCCAGTTCC, and M2-9: GTGGTGATGATGAGAACCACC TCGCTCAGACAGAATCACC. The second reverse primer, containing an EcoR1 cutting site was the same for all mutant proteins: ATAGCTATCCCAAGCTTAGTGATGGTGGTGATGATGAGAACC. Following PCR, the resulting fragments were column purified and digested with BamH1 and HindIII restriction enzymes, then column purified again. In parallel, the *surA* gene was digested with BamH1 and HindIII and purified by gel. The vector and fragment were ligated with DNA ligase for 6 hours at 16°C. The ligation product was transformed into JW0052 Δ*surA* strain as described above. The plasmid was purified and sequenced. Correct sequences were transformed into ER2566 strain and the proteins purified as described above.

2.9 Analysis of SurA Expression by Western Blot

One colony was picked from each mutant and grown overnight in 15mL of LB media at 37°C with shaking. The OD_{600} was taken for each culture. 13.5mL was taken from each culture and centrifuged at 5500 *x g* for 20 minutes. The cell pellet was re-suspended in 8mL of TRIS-sucrose solution (20% sucrose, 30mM TRIS, 1mM EDTA, pH 8.0) and incubated on ice with shaking for 10-20 minutes. The cell-containing solution was then centrifuged at 8000 *x g* for 20 minutes. The cell pellet was then re-suspended in a normalized volume of 5mM $MgSO_4$ based on the OD_{600} of each culture. The resulting supernatant was run on a 10% SDS-polyacrylamide gel for 50 min. The proteins were transferred to a nitrocellulose membrane at 20 V for 20 min. The membrane was blocked overnight with 1% BSA solution. Polyclonal anti-SurA antibody was used at a 1:4000 dilution in PBS and incubated for at least 2 hours. Anti-rabbit alkaline phosphatase antibody was used as the secondary antibody at a dilution of 1:3000 for 1.5 hours. The membrane was washed 3 times (PBS with 1:1000 dilution of Tween-20) after each application of the antibody. The membrane was developed with solutions NBT and BCIP and quenched by washing with DI water.

2.10 Limited Trypsin Digestion of SurA Mutants

Aliquots of each SurA mutant in the original dialysis buffer (20mM Sodium Phosphate, pH 7.5) were added to separate 1.5mL microcentrifuge tubes with approximately __ug of protein. Porcine trypsin was added to the SurA solution, corresponding to a trypsin:SurA mass ratio of 1:500. Immediately after the addition

of trypsin, a 15uL aliquot was removed and every 10 minutes afterwards through 60 minutes. To each aliquot removed, SDS-PAGE loading dye (containing 2% SDS and 10mM DTT) was added and the solution boiled for 5 minutes to quench the trypsin digestion. The samples were loaded onto either a 15% or 20% SDS-acrylamide gel and subjected to electrophoresis.

Chapter III: Results and Discussion

3.1 Importance of C-terminus for SurA Function

The goal of this study is to identify the critical residues responsible for the chaperone activity of SurA. As discussed above, although SurA contains four domains, the two PPIase domains are not vital for the *in vivo* activity of SurA. The C-terminal domain, however, has been shown to be critical for SurA function. As shown in Figure 3.1, the C-terminal domain is a long alpha-helix followed by a single beta-strand, which forms an anti-parallel beta-sheet with a beta-strand from the N-terminal domain. We hypothesize that this interaction is essential for the activity of SurA. To test this hypothesis, we created different constructs of truncated SurA, in which a small fraction of the sequence from the C-terminal end was systematically truncated. The activities of the resultant mutants were examined through introducing each protein into a *surA* gene knockout strain and examining the susceptibility level of the stain toward an antibiotic novobiocin. The minimum inhibitory concentration (MIC) of novobiocin for the mutants can indicate the relative activity of SurA. *E. coli* strain JW0052 (a Δ*surA* strain) containing plasmid encoded wild type SurA (WT) was used as the positive control. As a negative control, JW0052 containing the empty plasmid pMAL III was used.

SurA function is very sensitive to C-terminal truncations. A deletion of as short as 4 amino acids yielded a significant drop of activity, while deletion of the last 10 amino acids completely abolished the activity of SurA (Figure 3.1). As shown in Figure 2, the C-terminal beta-strand (red) closely packs against a N-terminal beta-

strand (green). When 10 residues of the C-terminal tail are removed, this interaction is lost (Figure 3.1, right), and this loss could destabilize the protein, preventing it from functioning properly. Additive sequential deletions from the C-terminal end seem to affect the MIC by decreasing the value as the number of deleted amino acids increases. From 2AA deletion to 4AA deletion, the MIC values decrease by ~2-fold with respect to WT. From 4AA deletion to 5AA deletion, the MIC values decrease further 1.5-fold. When 10 amino acids are deleted from the C-terminus, the MIC value is equal with that of Δ*surA* cells, indicating an inactive SurA protein.

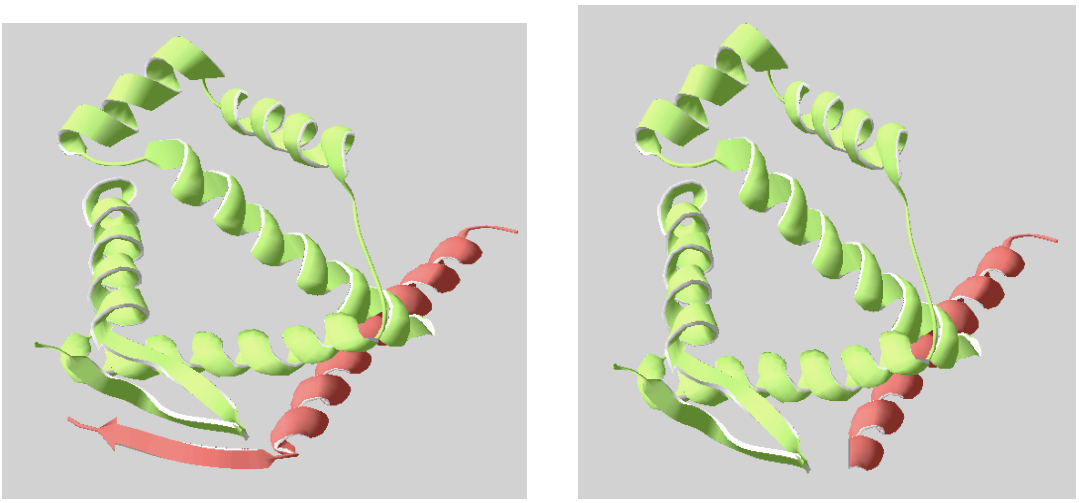

Figure 3.1 Structural images of N + C terminus of SurA (left) and N+C terminus with 10 amino acid deletion at C-terminus (right). N-terminus is shaded green and C-terminus is shaded red. Structure was generated from the Swiss Protein Data Bank, ID: 1M5Y [Bitto, E. and D. B. McKay (2002)].

Table 3.1 MIC of Truncated SurA mutants

Mutation	MIC of Novobiocin (ug/mL)
WT	1100
ΔsurA	100
20AA C-terminal deletion	100
11AA C-terminal deletion	100
10AA C-terminal deletion	100
5AA C-terminal deletion	250
4AA C-terminal deletion	350
2AA C-terminal deletion	256*

*Liquid LB was used where MIC of WT was 256 ug/mL novobiocin. All other experiments were conducted using the agar plate method.

3.2 Site-Directed Mutagenesis of Residues at the C-terminus of SurA

To further pinpoint key residues in the C-terminal domain of SurA, we conducted BLAST search using the sequence of SurA from *E. coli* to identify conserved residues in proteins present in distant Gram-negative species. Several conserved residues were identified and nine were chosen for site-directed mutagenesis at the C-terminus of the protein. The sequence alignment is shown below in Figure 3.2.

```
NW     TPKKKEIREKMFADKYEKRSKSYLERIRKSAMIEYR
ZM     QPNFQQIHNQLQEDRVNKRAIRYLRDLRRDAIIDYR
VV     SAMKNKAYRILFNRKFNEEVGAWMQELRAGAFVEII
BH     LIFSLQDNKQKSPQELEILSEKYLKELRQVARIKSS
EC  391AAQKDRAYRMLMNRKFSEEAASWMQEQRASAYVKIL426
                                :  .  *  *  :
```

Figure 3.2 Sequence alignment of *surA* gene across several distant gram-negative species. *Nitrobacter winogradskyi* (NW) *Zymomonas mobilis* (ZM)*Vibrio vulnificus* (VV) *Bartonella henselae* (BH) * indicates total conservation, : indicates conserved residue type, . indicates mostly conserved residue type.

We have created nine single mutation mutants in which the indicated amino acid was replaced by alanine, glycine, or serine. The identities and locations of the mutated residues are shown in Figure 3.3. JW0052 containing plasmid encoded wild-type SurA was used as the positive control. For the negative control, the *surA*-containing pMal III plasmid was modified to exclude the signal peptide sequence (cyto SurA) necessary to translocate the SurA protein into the periplasm. Without the signal peptide sequence, SurA remains in the cytoplasm and is not functional. A summary of the MIC values of novobiocin for some of mutants in liquid LB broth culture can be found in Table 3.2. The MIC values of novobiocin for the remaining mutants on LB agar media is presented in Table 3.3.

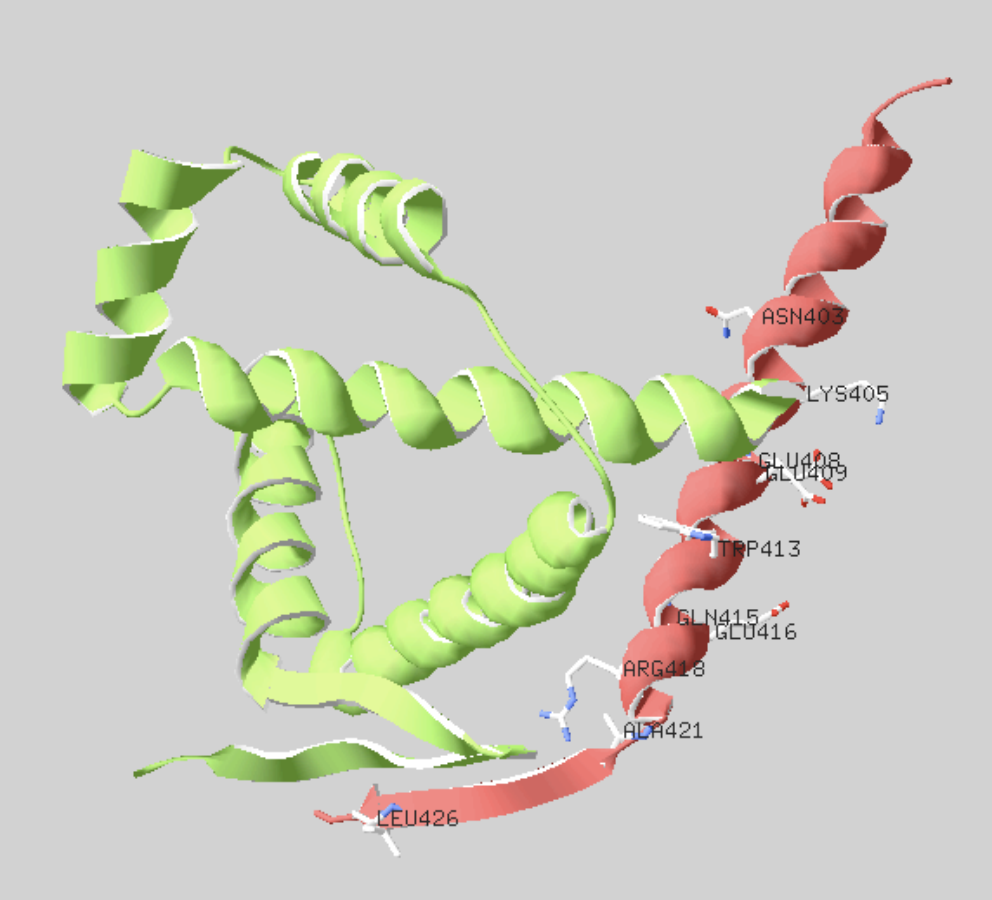

Figure 3.3 Structural image of SurA showing all residues on C-terminus that were mutated. Structure was generated from the Swiss Protein Data Bank, ID: 1M5Y [Bitto, E. and D. B. McKay (2002)].

Table 3.2 MIC values of novobiocin for mutants at C-terminus of SurA in liquid LB broth .

Mutation	MIC of Novobiocin (ug/mL)
WT	256
Cytoplasmic SurA	64
W(413):G	200
R(418):G	200
W(413):G R(418):G	160
I(425):S	256
E(416):A	256
K(405):A	256
E(408):A	128
N(403):A	256
Q(415):A	256
3AA C-terminal deletion + I(425):S V(423):L A(421):W	128

Novobiocin was added to the liquid culture at concentrations ranging from 16-256 μg/mL at increments of 2-fold differences in concentration. The strain containing cytoplasmic SurA showed 16-fold decrease in MIC from WT, due to the absence of SurA in the periplasm. This shows that lack of SurA activity will render cells much more susceptible to bulky hydrophobic antibiotics than cells with

properly functioning SurA. The only single mutation variant to show a consistent decrease in MIC of at least 2-fold from the WT was E(408):A. MIC values for all other mutants averaged out to be less than a 2-fold difference. The SurA construct containing 3 residue deletion at the C-terminus plus I(425):S V(423):L A(421):W triple mutation was a result of a unintended single-base deletion at the G(1263) position and subsequent frame-shift mutation. Similar to the 4AA deletion, which showed a difference of 2-fold decrease in MIC from WT (Table 3.1), this mutant with 3 additional single mutations and a 3AA deletion resulted in the same 2-fold decrease in MIC. This result shows that SurA is a very robust protein, capable of functioning with many substitutions at the C-terminus. The importance of residues at the C-terminus has been established by the deletion mutants from Table 3.1, but the specific residue and even the class seems to be unimportant for function. The shift from alanine, a small hydrophobic residue to tryptophan, a bulky aromatic residue, coupled with the shift from isoleucine, a branched hydrophobic residue, to serine, a hydroxyl-containing polar residue had little or no effect on the function of SurA. It is unlikely that the C-terminal tail takes any part in binding the targeted unfolded proteins destined for the outer membrane, but it could play a vital role by associating with the N-terminus for the overall stability of the protein.

After many experiments and inconsistent results with the broth MIC method, agar plates were used for verifying the MIC values of mutants. The agar plate method provided easier determination of MIC by simply identifying the presence of visible colonies, rather than relying on optical density values, which were often inconsistent. However, as a result, there is a disparity in the numbers for

the MIC. The MIC values from the agar plate method are much higher than those found when the LB broth method was used. There is no clear reasoning for this. It is possible that the agar has some antagonistic effect on the drug itself and its uptake into cells. It is also possible that some of the novobiocin is broken down during preparation of the agar plates. The drug must be added when the agar media is well above the gelling temperature of the agar (32-45°C). At a temperature above 45°C, the novobiocin could breakdown. Even though the reasoning for differences between the two methods is not well understood, the agar plate method still provided much more consistent data and was used for that reason.

The relative MIC differences between JW0052 containing wild type, cytoplasmic, and other SurA mutants are fairly consistent. The E(408):A mutant did not grow at a novobiocin concentration of a third the wild type MIC value in both the agar and broth methods. The Δ*surA* strain was not tested using the broth method, but could be compared with cytoplasmic SurA from that data set because they have the same end result, lack of functioning SurA. These two have similar differences in MIC, approximately 8-fold, when compared with WT. Thus, the important value for MIC is not the raw number, but its relation to the controls.

Table 3.3. MIC values of novobiocin for mutants at C-terminus of SurA on LB agar plates

Mutation	MIC of Novobiocin (ug/mL)
WT SurA	1100
Δ*surA*	100
E(408):A	350
S(407):A E(408):A	400
E(408):A E(409):A	400
E(409):A	450
S(407):A	500

Several more single and double mutation variants were tested for MIC using the agar plate method. None of these achieved a reduction in MIC, 2-fold or greater against WT, except the E(408):A mutant. One would expect that the two double mutation variants, which include the E(408):A mutation, would be equal or lower in MIC, but the averages came out higher. Due to the nature of the method and lack of resolution, it can be challenging sometimes to assess whether or not colonies have formed. Only increments of 50 μg/mL are used, so the MIC could lie between the two values. If more trials were conducted with these three mutants, the averages would likely converge. The important conclusion from this is that the adjacent amino acids on either side of E(408) have a minimal effect on the function of SurA. Furthermore, additive mutations do not necessarily equal the sum of the effects of each individual mutation.

3.3 Characterization of $SurA_{E408A}$ and WT SurA

Following MIC tests, one of the mutants (E408:A) was chosen for structural characterization based on fluorescence emission and circular dichroism (CD) spectra. $SurA_{E408A}$ was chosen because of all the single mutations attempted, it seemed to have the most significant impact on the MIC of novobiocin in several independent experiments.

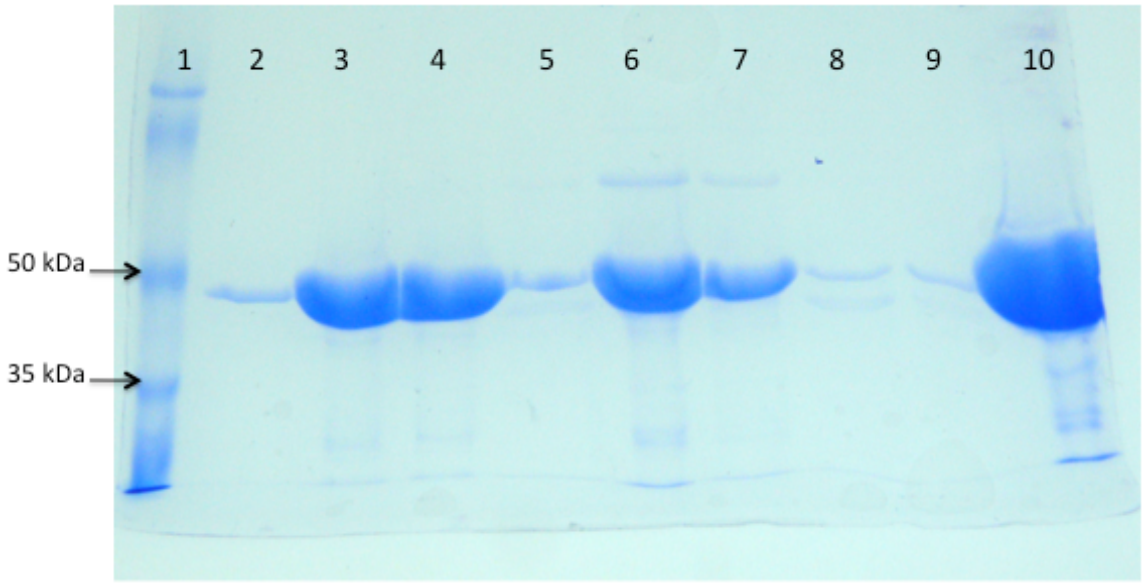

Figure 3.4 SDS-PAGE gel of purified WT SurA and E(408):A SurA. Lane 1: Molecular Weight Ladder, Lane 5: E(408):A SurA Wash, Lane 6: E(408):A SurA Fraction 1, Lane 7: E(408):A SurA Fraction 2, Lane 8: WT SurA Wash, Lane 9: WT SurA Fraction 1 , Lane 10: WT SurA Fraction 2

$SurAE_{408A}$ was expressed and purified. Figure 3.4 showed the representative SDS-PAGE . $SureA_{E408A}$ expressed as a similar level as the wild type SurA and could be purified with similar efficiency, indicating that the single point mutation did not have a drastic effect on the stability of the overall protein structure. E(408) is on the C-terminal alpha-helix and the acidic side-chain faces away from the interface with the N-terminus and is not closely associated with any other region of the protein. It is reasonable to postulate that the mutation to alanine will not severely affect the association between the N and C-termini, thereby retaining the structural integrity of the core module. Using the purified protein, we have collected its CD spectra (Figure 3.5 and 3.6) as well as intrinsic fluorescence spectrum (Figure 3.7). The concentrations of both proteins, following dialysis, were approximately 50 μM, as measured by the absorbance at 280 nm.

Following purification and dialysis, the secondary structure of both $SurA_{E408A}$ and WT SurA was analyzed by circular dichroism (CD). Circular dichroism is the measurement of the difference in absorption between right and left circular-polarized light. Secondary structure circular dichroism is used to probe the protein for alpha-helices and beta-sheets from 190-260 nm with high reliability, where the chromophore is the peptide bond [Johnson, W. C. (1990)]. A protein composed primarily of alpha-helices will show a characteristic double trough. Those that are primarily anti-parallel beta-sheet will show a single trough and a large peak around 195 nm [Kelly, S. M., T. J. Jess, et al. (2005)]. Both SurA structures show the characteristic alpha-helix double trough at ~ 222 and 210 nm (Figure 3.5). There is a small change of peak intensity between the plots of wild type and mutant SurA. No

peak shifts are observed in any spectra. From this CD plot, it can be concluded that there is still high helical contents in the structure of $SurA_{E408A}$ with little deviation from WT SurA.

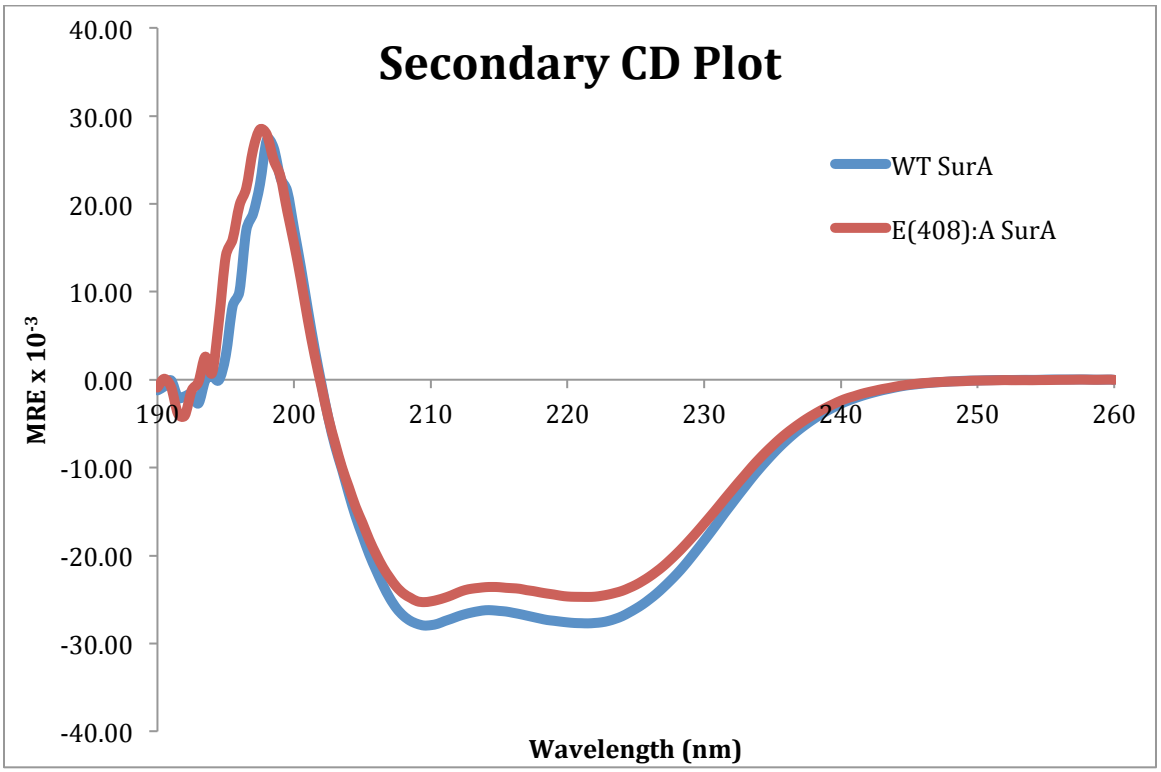

Figure 3.5. Circular Dichroism plot of SurA protein samples $SurA_{E408A}$ and WT SurA. MRE = Mean Residue Ellipticity (deg*cm^2/dmol).

Next, protein samples $SurA_{E408A}$ and WT SurA were analyzed for differences in tertiary structure using CD. The wavelength range is different for tertiary structure because the chromophores are aromatic side chains and disulfide bonds rather than peptide bonds for secondary CD. Due to the complex nature of the near-

UV CD spectra, only relative changes between a protein and mutants of that protein are reliable [Kelly, S. M., T. J. Jess, et al. (2005)]. In spite of its limitations, tertiary CD studies have yielded very important information regarding the structural changes of proteins subjected to site-directed mutagenesis [Woody, A. Y. M. and R. W. Woody (2003), Boxer, D. H., H. Zhang, et al. (2004)]. While the two plots were very similar, there is a small change of peak shape at around 290 nm, indicating a small change of the tertiary structure of $SurA_{E408A}$ (Figure 3.6).

.

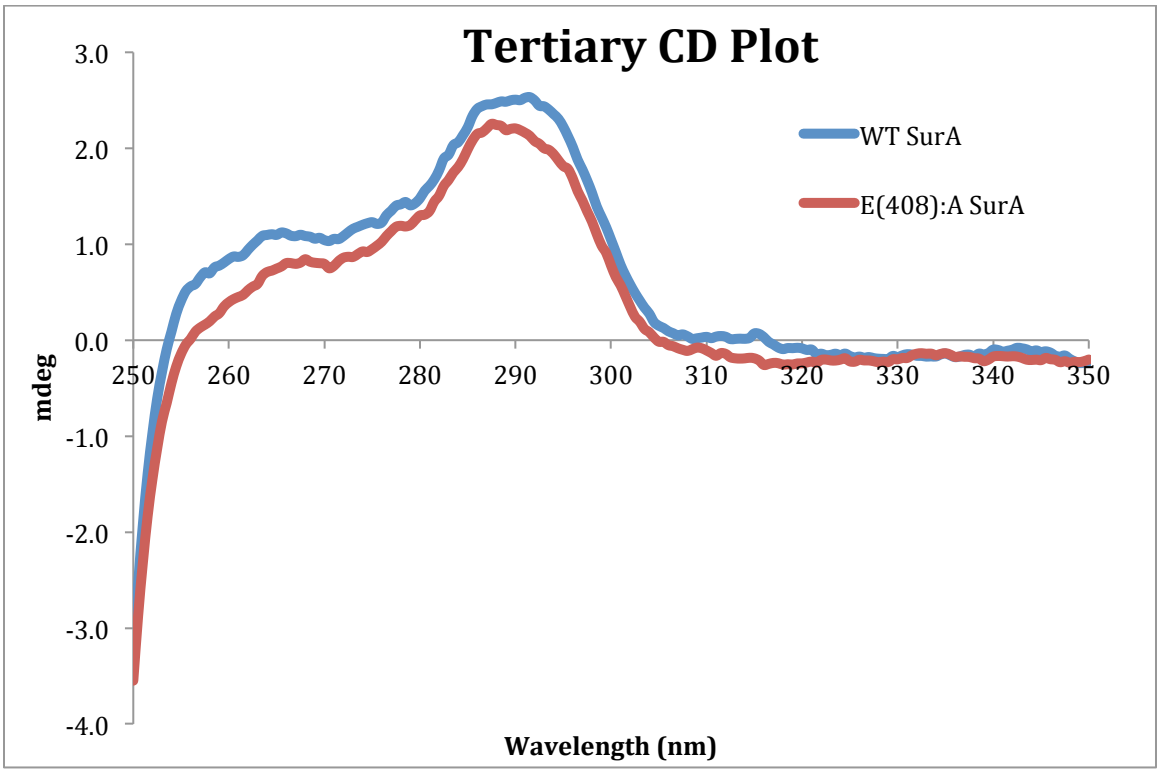

Figure 3.6. Circular Dichroism on protein samples $SurA_{E408A}$ and WT SurA

The tertiary structure of WT SurA and $SurA_{E408A}$ were further probed using fluorescence spectroscopy. Fluorescence emission spectroscopy is a popular tool to reveal relative changes in protein tertiary structure. The chromophores in this method are the intrinsic tryptophan, tyrosine, and phenylalanine residues of the protein. Of the three, tryptophan contributes most to the fluorescence scan due to its high quantum yield and quenching of tyrosine residues [Teale (1960)]. SurA has five intrinsic tryptophan residues, making it possible to analyze small changes in the tertiary structure. Fluorescence emission spectra are complicated to interpret, but the most important aspect of a scan is the position and intensity of the peak. This indicates the environment around the tryptophan residues or the relative exposure of tryptophan to the solvent . Tryptophan residues, which are buried in the protein are blue-shifted relative to those which are exposed [Vivian, J. T. and P. R. Callis (2001)]. However, scans of proteins containing more than one tryptophan are extremely difficult to interpret because the peak will be the sum of all of the signals from tryptophan. While some tryptophans could become buried, the others become exposed. On the other hand, if the protein is completely destabilized, this will be evident in the emission scan. Intensity of the peak has little to do with the position and orientation of tryptophan residues, but could be due to differences in concentration, or a combination of factors that are not fully understood. In the plot (Figure 3.7), there is little difference in peak position, but a small change of peak intensity. The maximum emission peaks for WT SurA and $SurA_{E408A}$ were 342.0 nm and 341.5 nm, respectively. This is expected because the range of tryptophan emission is between 308nm for fully buried residues, and 355nm for fully exposed

residues [Vivian, J. T. and P. R. Callis (2001)]. The lack of peak shift between the two proteins suggested that there was no large scale change of the global structure.

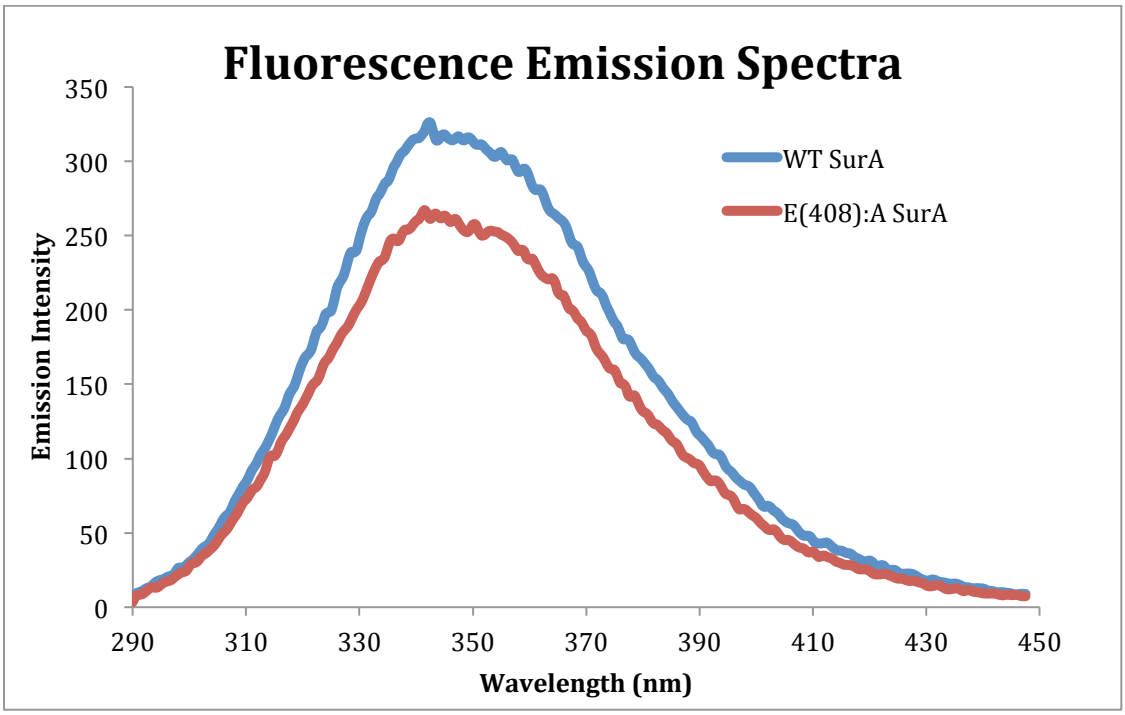

Figure 3.7. Fluorescence emission spectra for E(408):A SurA and WT SurA over the range 290-450 nm. Excitation wavelength corresponding to maximum tryptophan absorbance was 280nm.

Next, to examine the stability of the $SurA_{E408A}$ and WT SurA protein structures, a temperature melt experiment was performed. Using the CD spectropolarimeter, readings were taken at 222 nm as the temperature of the solution was heated from 4°C to 98°C (Figure 3.8). The temperature melting curve provides information about the stability of the protein. The slope of the curve indicates how well packed the protein is and how cooperative its interactions are.

The T_m or melting temperature of the protein represents the temperature where 50% of the protein is unfolded and is indicative of its stability. Proteins with a higher melting temperature are more stable than those with a lower melting temperature. Both protein samples have approximately the same melting temperature of 55°C. The difference between the two protein samples is the slope of the curve (Figure 3.9). $SurA_{E408A}$ has a slightly sharper slope than WT SurA. The sharper slope suggests greater cooperativity of protein folding in the mutant. However, the change is slight and more evidence would be needed to draw a definitive conclusion on the relative stability of these two proteins.

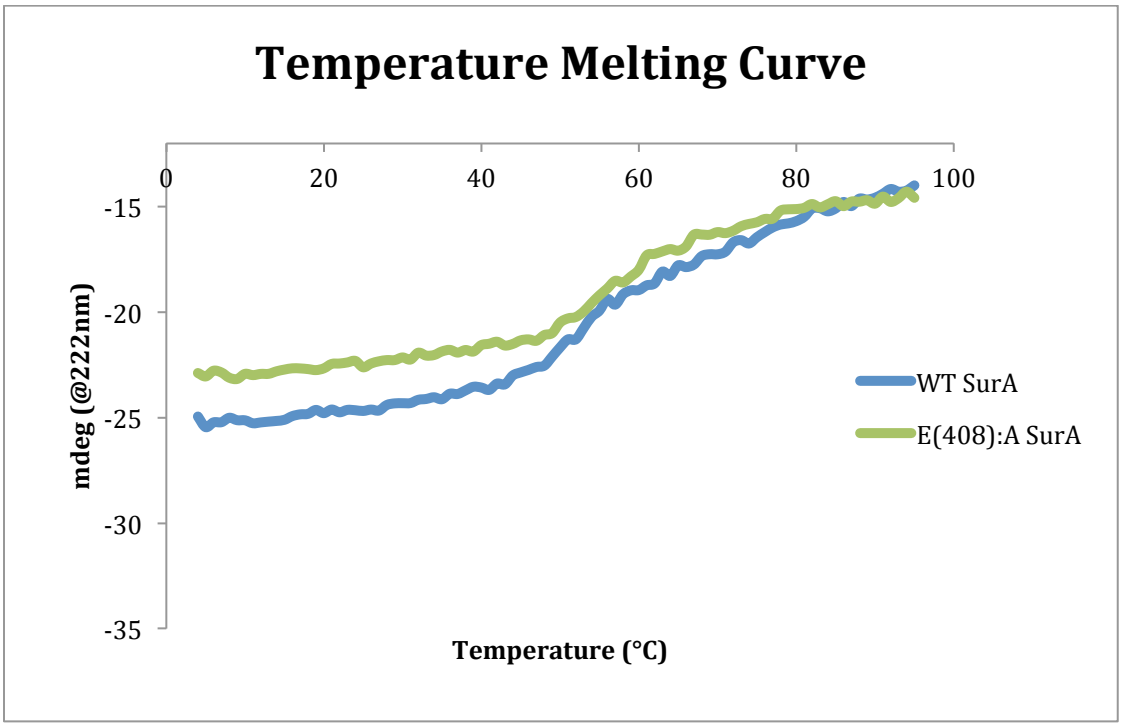

Figure 3.8 CD readings at 222 nm at a rate of 2°C/min.

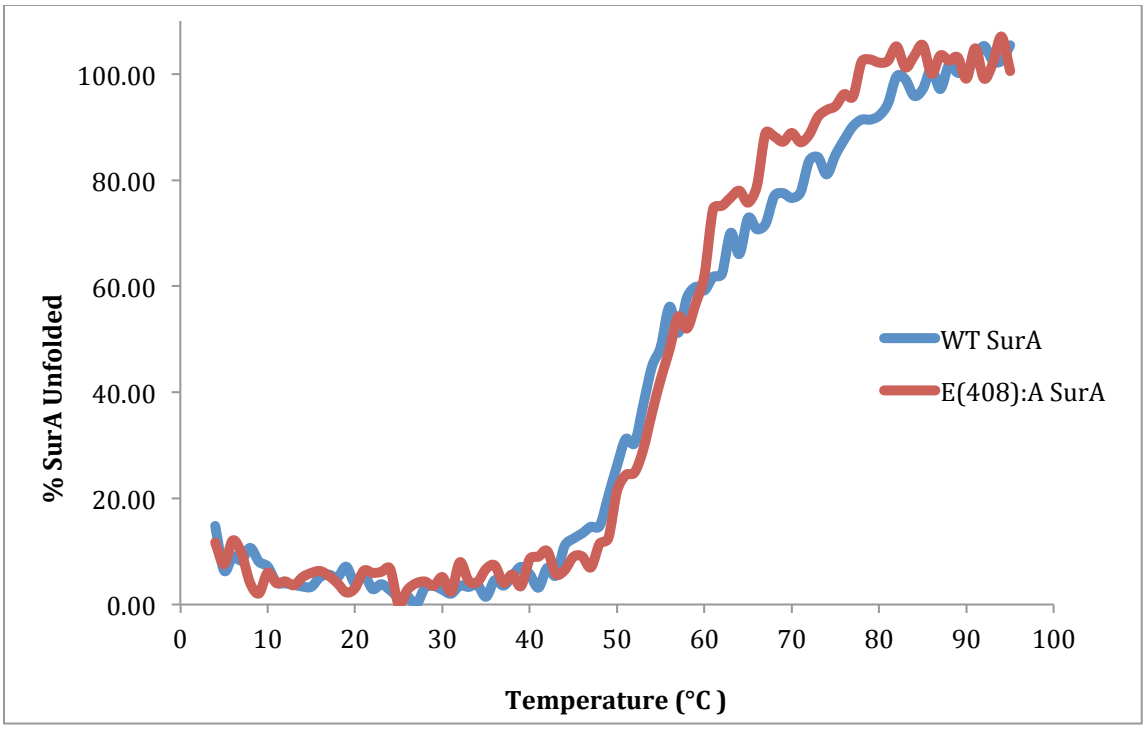

Figure 3.9 This plot shows the relative percentage of WT SurA and E(408):A SurA unfolded during the temperature melt.

Next, WT SurA and $SurA_{E408A}$ proteins were analyzed for function by testing the binding efficiencies to a denatured outer membrane porin (OmpF). This was accomplished by employing fluorescence polarization. Fluorescence polarization works by detecting changes in the tumbling of fluorescent molecules in solution. If a molecule binds another, then the tumbling will slow down and this is detected by the fluorometer [Jameson, D. M. and S. E. Seifried (1999)]. Thus, it is a valuable method for observing binding between two molecules. OmpF was purified from *E. coli*, then denatured in 6 M urea. WT SurA and $SurA_{E408A}$ were labeled with the fluorophore, fluorescein isothiocyanate (FITC), which covalently links to the free

amine on the side chain of lysine residues. Following dialysis to remove free FITC in the reaction mixture, SurA was added into a cuvette, followed by titration using OmpF to observe binding (Figure 3.10).

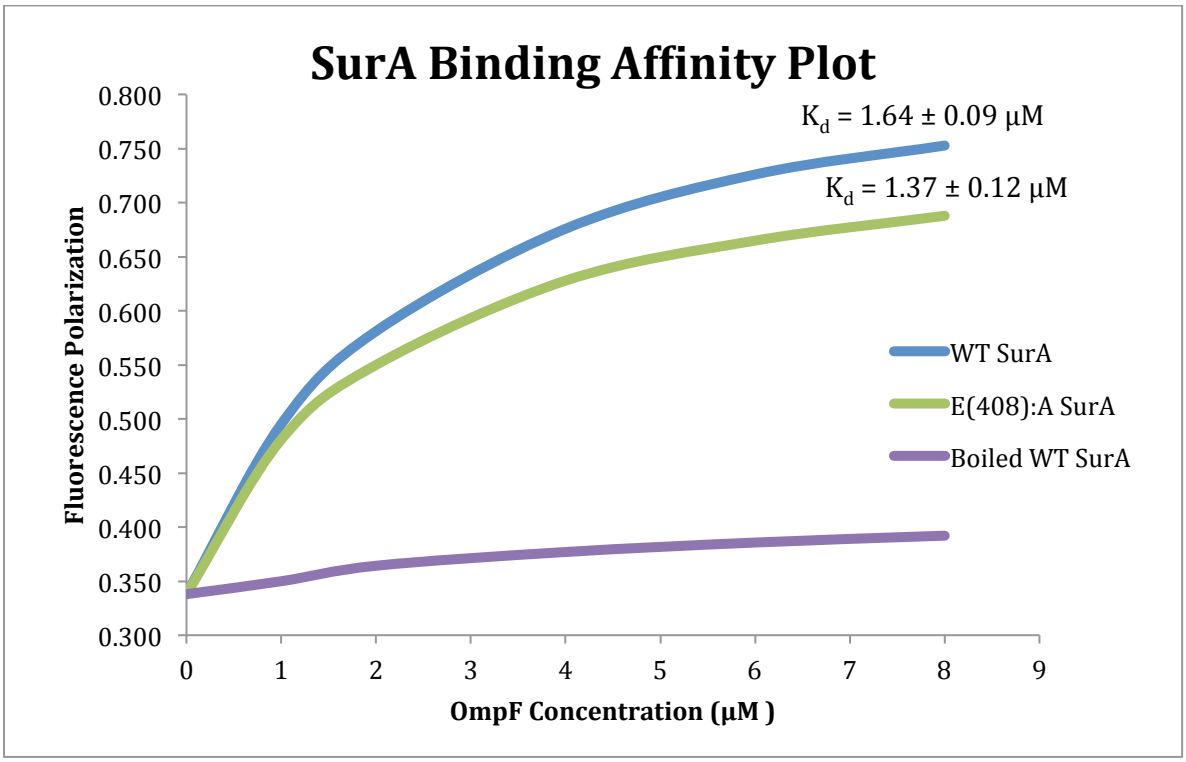

Figure 3.10 This plot shows changes in fluorescence polarization with increasing concentrations of OmpF. WT SurA protein was boiled for 5 minutes to show that when it is unfolded, it binds very poorly to OmpF.

The result from the binding assay indicates little difference in the ability of $SurA_{E408A}$ and WT SurA to bind OmpF. The K_d values were calculated for both samples and although the $SurA_{E408A}$ sample gave a lower K_d value, indicative of

tighter binding, the difference is minimal. Thus, the mutation E(408):A in SurA is not detrimental to the chaperone function of SurA, as evidenced by its ability to bind unfolded OmpF with the same affinity as WT SurA.

In conclusion, from the evidence of CD spectra and fluorescence emission spectra, the structure of $SurA_{E408A}$ is not significantly different from WT SurA. The peaks of both samples in all spectra shown had strong overlap with little or no changes in the position of the peak. The binding affinities with OmpF as measured *in vitro* were also very similar. What caused the observed decrease in MIC values of novobiocin remains a mystery. More in-depth structural analyses would be necessary to determine the cause of the partial loss of activity in $SurA_{E408A}$.

3.4 Importance of Hydrophobic Interactions for SurA Function

Hydrophobic interactions can be a very important factor for a protein with a wide range of substrate specificity. Aromatic residues have bulky hydrophobic moieties that could be vital for the hydrophobic interactions necessary for low specificity binding. Only the C and N-termini are vital for SurA function, so the aromatic residues at these two domains were the focus of this study. To test whether the aromatic residues at the C and N-termini are vital for the function of SurA, a large selection of them were mutated to alanine. Four double mutants and five triple mutants were also constructed to further elucidate the role of aromatic residues in SurA function.

Table 3.4. MIC values for aromatic residues in the N-terminus and C-terminus mutated to alanine. Concentration of novobiocin on LB agar plates ranged from 100μg/mL to 500μg/mL.

Mutation	MIC of novobiocin (μg/mL)
W(413):A	>500
Y(398):A	>500
Y(128):A	>500
Y(120):A	>500
Y(422):A	>500
Y(125):A	>500
F(406):A	500
Y(398):A Y(422):A	>500
Y(120):A Y(128):A	>500
Y(120):A Y(125):A	>500
Y(125):A Y(128):A	>500
Y(398):A Y(422):A Y(120):A	>500
Y(398):A Y(422):A Y(125):A	150
Y(398):A Y(422):A Y(128):A	150
Y(398):A Y(422):A F(406):A	>500
Y(120):A Y(125):A Y(128):A	>500

From the MIC data gathered here, it seems that these mutations have little effect on the function of SurA. The only single mutation that caused any decrease in MIC was F(406):A. All of the double mutants were found to be active at concentrations of novobiocin tested. At least three alanine substitutions were required to inhibit SurA function significantly and in both cases, Y(398) and Y(422)

must be mutated. It is unclear why these two positions, coupled with either Y(125) or Y(128) are vital for proper function. It is surprising that the iterative mutations leading up to the triple mutant showed no loss in activity. It is possible that because these three positions are in three different corners of the core module, mutations will de-stabilize the entire core module, cooperatively, to the point where it is no longer functional. The Y(398, 422, 120):A mutant did not exhibit this phenotype. This is likely due to the orientation of the aromatic side-chain of Y(120):A. The side chain in Y(120) is exposed, and faces away from the core-module, while those of both Y(125) and Y(128) are buried and face into the core module. Disrupting buried, bulky, hydrophobic residues has a much larger effect on the stabilization of the protein than disrupting those which are exposed. With these results, we show that additive mutations on the aromatic residues of SurA have a deleterious effect for the protein. Whether this is a result of decreased OMP binding or de-stabilization of the protein cannot be answered with the MIC assay alone.

3.5 Sequence tolerance of C-terminus of SurA

The data obtained from site-directed mutagenesis has revealed that SurA is a very robust protein, and can accommodate many different single and double mutations at the C-terminus without compromising the function. However, when the C-terminus is truncated by 10 amino acids, the function is completely lost (Table 3.1). The focus of this final study was to determine if the identity or class of the last 10 amino acids were important for function.

Toward this end, a *surA* vector (denoted EC-1) was constructed in which the gene sequence corresponding to the last 10 residues was removed, while two restriction sites (NheI and EcoR1) were introduced to allow for a 30nt random library to be inserted at the C-terminus. EC-1, when transformed into Δ*surA* cells showed the same activity profile as the 10 AA truncated protein with novobiocin treatment (Table 3.5). Due to the large difference between WT *surA* activity and EC-1 activity in the presence of novobiocin, a wide range of mutants with a random sequence at the C-terminus could be identified.

Sequences that could restore the activity of SurAΔ419-428 were identified using a novobiocin assay. After ligation, the library of plasmids were transformed into JW0052 and then plated on novobiocin-containing agar media (150 μg/mL). We determined at this concentration JW0052 or JW0052 containing EC1 did not grow. Therefore, colonies that grew on the plate must have contained partially active SurA. To test the relative frequency of mutants that could recover SurA function, the number of colonies that grew with novobiocin selection pressure were divided by the number of colonies that grew without selection pressure. The average frequency was ~ 1.5% from three plates in each testing group. Thus, out of 1.0×10^{13} possible sequences, an estimated 1.5×10^{11} sequences will partially restore SurA activity.

Colonies that grew on novobiocin selecting plates were tested for MIC by first purifying the plasmid, and then picking three colonies following re-transformation. This was done to ensure that the resistance to novobiocin was due to the plasmid

and not a genomic mutation. In total, 26 colonies were chosen and analyzed, with a wide range of activity in the presence of novobiocin and SDS-EDTA (Table 3.5). Among the 26 mutants, 14 exhibited activity (MIC ≥ 900 μg/mL) close to or better than WT for the novobiocin MIC assay, and 11 of them were viable when subjected to concentrations of 5% SDS (1mM EDTA) or greater. The two assays differed somewhat in the selectivity relative to WT. This is not surprising, as these two chemicals affect the bacteria in very different ways. SDS is a harsh detergent and works by perturbing and dissolving the outer membrane directly. EDTA is added to chelate divalent cations, Mg^{2+} and Ca^{2+}, further destabilizing the outer membrane. Novobiocin, on the other hand, must traverse both membranes before it can act as an inibitor of DNA gyrase in the cytoplasm.

There seemed to be no pattern in the sequences of the mutants chosen. Only at position 5 is there any sequence homology between the mutants. At this position in the most active mutants (MIC ≥ 1000 μg/mL novobiocin) there is either a valine, isoleucine, or methionine. WT SurA has a valine at this position, but both isoleucine and methionine are very good substitutions for valine. The role of valine at this position is not known, as it was never mutated in any of the previous work. However, it is not likely that it is vital for function because the mutant M4-19 has a cysteine substituted there and still exhibits good activity in the presence of both novobiocin and SDS-EDTA. Thus, we conclude that the presence of the last 10 amino acids is vital for function, but the sequence is not necessarily very important.

Table 3.5 Sequences of last 10 amino acids at C-terminus of *surA* and MIC values of all mutants that were selected. Colonies were tested at concentrations of SDS from 1.5% to 5% on LB agar media, all containing 1mM EDTA. Higher concentrations were not possible due to solubility issues. The range for novobiocin (nov.) concentrations was 500 μg/mL to 1200 μg/mL on LB agar media. The (*) indicates a stop codon.

Mutant	**MIC**		**Sequence of Last 10 AA**									
	Nov.	SDS+EDTA										
WT	**1100**	>5	**A**	**S**	**A**	**Y**	**V**	**K**	**I**	**L**	**S**	**N**
M2-9	1200	>5	**A**	**C**	**Q**	**L**	**V**	**I**	**L**	**S**	**E**	**R**
M4-1	1200	>5	**L**	**G**	**T**	**T**	**V**	**I**	**M**	**P**	**P**	*
M4-8	1200	5	**M**	**T**	**V**	**Y**	**I**	**R**	**L**	**T**	**H**	**S**
M4-15	1150	>5	**F**	**N**	**T**	**V**	**V**	**V**	**Y**	**L**	**S**	**S**
M4-7	1100	4	**V**	**S**	**S**	**K**	**V**	**S**	**L**	**C**	**I**	**R**
M4-9	1100	4	**T**	**I**	**C**	**H**	**V**	**A**	**I**	**P**	**R**	**G**
M4-10	1100	5	**Q**	**K**	**A**	**R**	**M**	**M**	**G**	**G**	**G**	**G**
M4-13	1100	>5	**A**	**L**	**A**	**W**	**V**	**I**	**Y**	**F**	**C**	**F**
M2-1	1100	4	**A**	**V**	**G**	**R**	**V**	**A**	**T**	**P**	**S**	**M**
M4-11	1000	5	**D**	**K**	**T**	**L**	**I**	**C**	**Y**	**E**	**H**	**C**
M1-1	1000	5	**L**	**K**	**A**	**Y**	**M**	**F**	**L**	**S**	**V**	**L**
M1-8	1000	3	**Q**	**H**	**I**	**T**	**V**	**R**	**V**	**A**	**G**	**G**
M2-5	1000	4	**A**	**R**	**I**	**F**	**V**	**R**	**L**	**F**	**G**	**L**
M4-19	950	>5	**F**	**F**	**C**	**R**	**C**	**F**	**L**	**T**	**S**	**D**
M3-5	900	2.5	**V**	**K**	**S**	**R**	**A**	**T**	**V**	**G**	**S**	**Q**
M4-20	900	>5	**L**	**L**	**A**	**V**	**V**	**I**	**M**	**S**	**V**	**Q**
M3-1	875	4	**I**	**R**	**W**	**L**	**K**	**T**	**V**	**I**	**I**	**R**
M3-6	875	2.5	**Q**	**P**	**A**	**H**	**V**	**W**	**V**	**V**	**L**	**S**
M4-12	750	2.5	**A**	**R**	**V**	**V**	**G**	**V**	**V**	**P**	**M**	**E**
M2-4	650	4	**I**	**R**	**G**	**P**	**V**	**V**	**R**	**V**	**V**	**Q**
M2-10	650	3	**V**	**C**	**A**	**R**	**V**	**R**	*			
M2-12	650	4	**L**	**S**	**T**	**H**	**I**	**S**	**V**	**V**	**S**	**V**
M4-21	650	>5	**A**	**C**	**A**	**T**	**V**	**V**	**W**	**G**	**G**	**S**
M2-8	600	2.5	**R**	**R**	**A**	**W**	**F**	**F**	**L**	**C**	**H**	**G**
M2-7	600	4	**D**	**L**	**L**	**R**	**R**	**Y**	**V**	**S**	**S**	**S**
AK-1	500	1.5	**A**	**I**	**T**	**L**	**L**	**L**	**G**	**C**	**S**	**S**

3.5.1 Relative SurA Expression Among Mutants

In order to rule out protein expression levels as a source for the differences in activity among the mutants, each was tested for expression of SurA. The periplasmic fraction was isolated from overnight cultures and run on SDS-PAGE before transferred to a membrane for Western blot analysis with an anti-SurA antibody (Fig. 3.11a-3.11d). The results indicated normal expression levels for nearly all mutants. A few, however, did show low expression of SurA; in particular, AK-1, M1-8, M4-8 (Fig. 3.11d), and M2-8(Fig. 3.11b). A low expression level does not necessarily confer low activity. For instance, M4-8 is among three mutants with activity better than WT in the novobiocin MIC assay. On the other hand, AK-1 has the lowest activity of all mutants for both MIC assays. It should also be noted that the template, EC-1, has expression levels comparable to WT and has no activity. Thus, it can be concluded that for most mutants full-length SurA existed in the periplasm at levels that were sufficient for normal function. Additional SurA did not further increase the observed SurA activity.

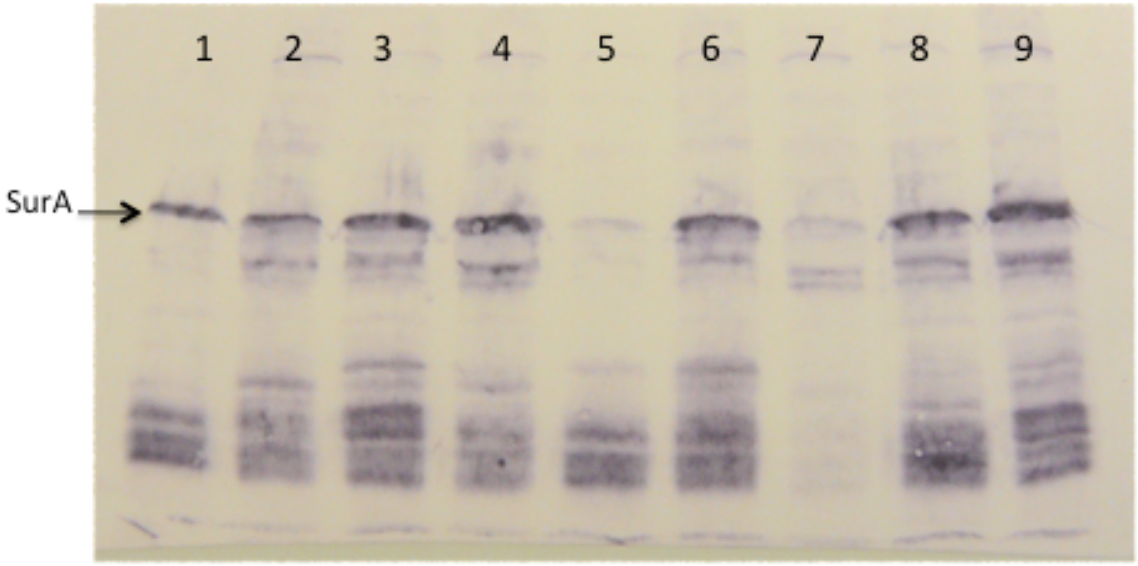

Figure 3.11a Western Blot using anti-SurA antibody. From left, Lane 1: WT SurA, Lane 2: M3-1 SurA, Lane 3: M2-1 SurA, Lane 4: M2-4 SurA, Lane 5: M2-5 SurA(re-tested Fig. 3.11d) , Lane 6: M1-1 SurA, Lane 7: M1-8 SurA (re-tested Fig. 3.11d), Lane 8: M3-6 SurA, Lane 9: M3-5 SurA.

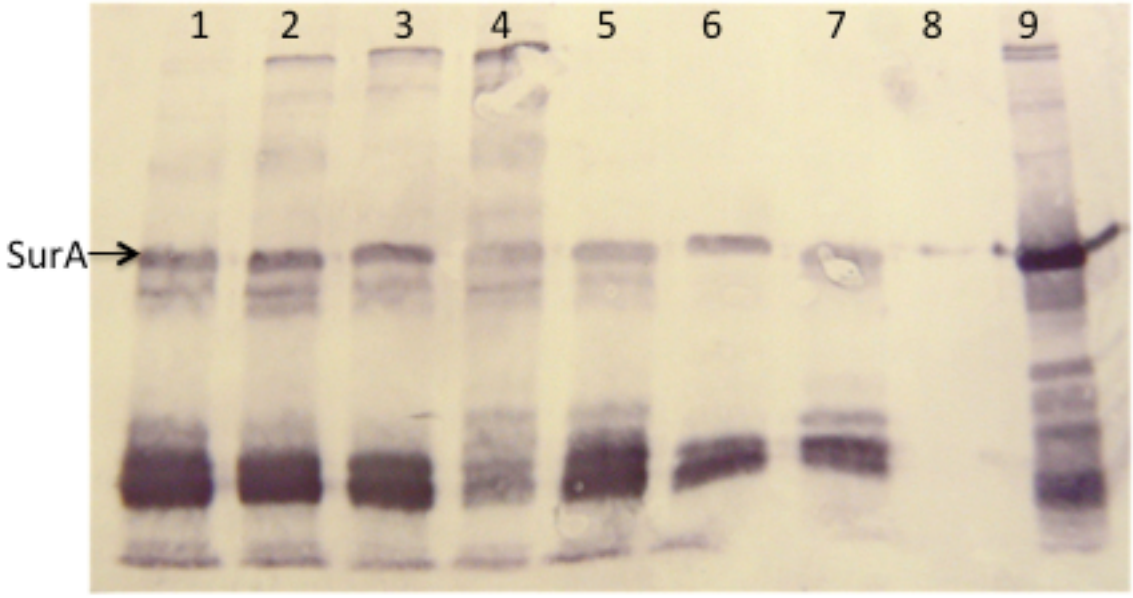

Figure 3.11b Western Blot using anti-SurA antibody. From left, Lane 1: M2-12 SurA, Lane 2: M2-10 SurA, Lane 3: M2-9 SurA, Lane 4: M2-8 SurA, Lane 5: M2-7 SurA, Lane 6: EC-1 SurA, Lane 7: WT SurA, Lane 8: empty, Lane 9: Purified SurA.

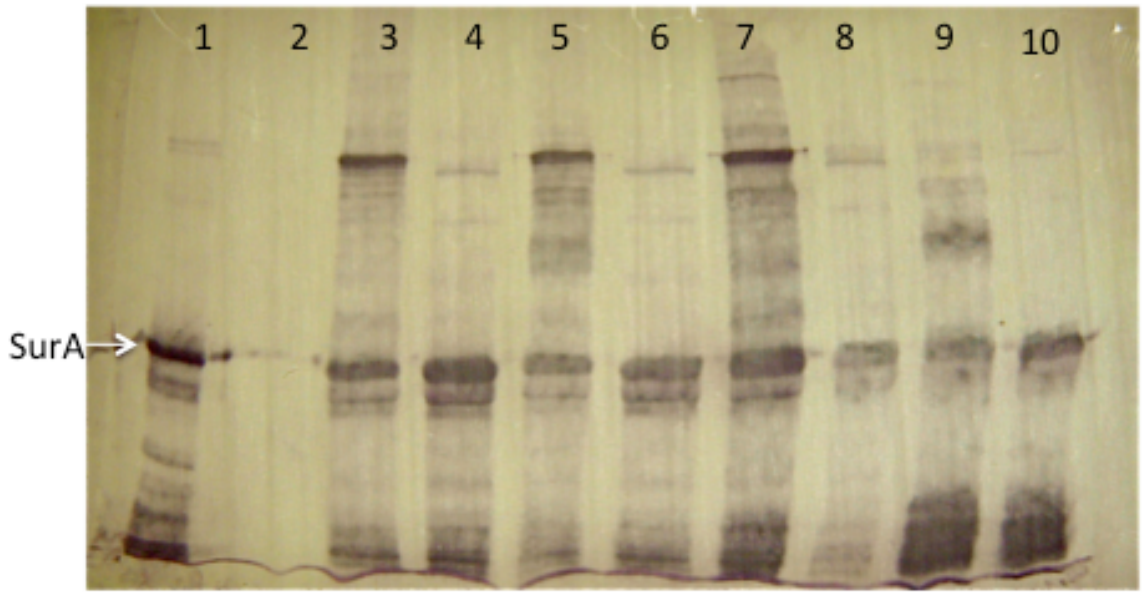

Figure 3.11c. Western Blot using anti-SurA antibody. From left, Lane 1: WT SurA, Lane 2: ΔsurA, Lane 3: M4-21 SurA, Lane 4: M4-20 SurA, Lane 5: M4-19 SurA, Lane 6: M4-15 SurA, Lane 7: M4-13 SurA, Lane 8: M4-12 SurA Lane 9: M4-11 SurA. Lane 10 : M4-10

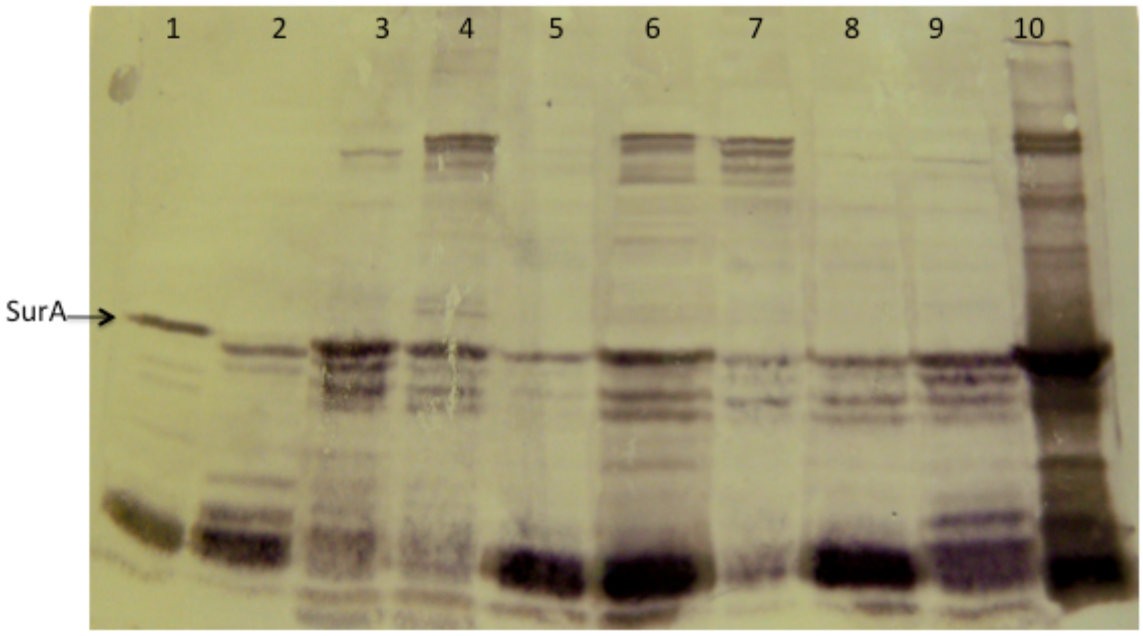

Figure 3.11d. Western Blot using anti-SurA antibody. From left, Lane 1: WT SurA, Lane 2: EC-1 SurA, Lane 3: M4-1 SurA, Lane 4: M4-7 SurA, Lane 5: M4-8 SurA, Lane

6: M4-9 SurA, Lane 7: AK-1 SurA, Lane 8: M1-8 SurA, Lane 9: M2-5 SurA. Lane 10 : Purified SurA

3.5.2 Characterization of $SurA_{M2-1}$, $SurA_{M2-7}$, and $SurA_{M2-9}$

SurA from the three mutants M2-1, M2-7, and M2-9 were chosen for protein characterization based on their activity and relative expression. All three had comparable expression levels: two (M2-1 and M2-9) showed high activity and one (M2-7) showed modest activity in the MIC assays. All three were expressed and purified, as described above, from the periplasm. Figures 3.12 and 3.13 show bands corresponding to SurA fractions of all three proteins. $SurA_{M2-9}$ has a band on the gel (Fig. 3.12) corresponding to a dimer. This is due to a cysteine introduced as part of the random peptide at the C-terminus. SurA has no intrinsic cysteine residues, which explains why the WT SurA shows no dimer formation. Two SurA monomers will dimerize via an S-S linkage between two cysteine residues in the oxidizing environment of the periplasm. This dimer formation is also seen on the western blot of M2-9 (Fig. 3.11b), as well as with other mutants that have cysteine additions at the C-terminus.

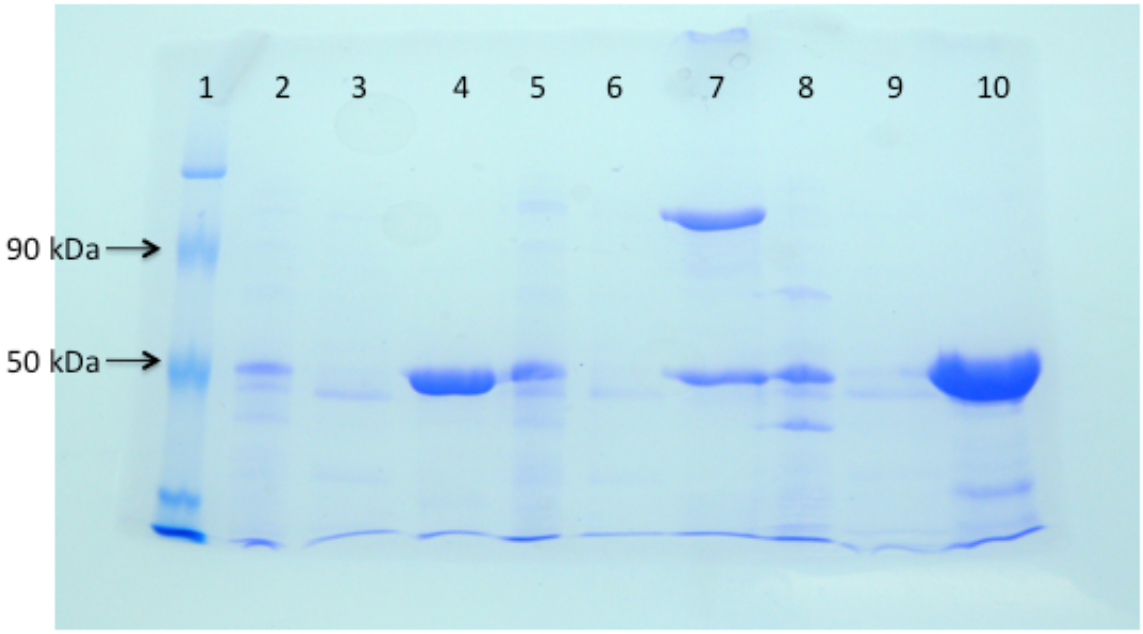

Figure 3.12. SDS-PAGE gel of purified WT SurA , M2-1 SurA, and M2-9 SurA. From left, Lane 1: Molecular Weight Ladder, Lane 2: Periplasmic fluid from M2-1 SurA, Lane 3: M2-1 SurA Wash, Lane 4: M2-1 SurA Fraction 3, Lane 5: Periplasmic fluid from M2-9 SurA, Lane 6: M2-9 SurA Wash, Lane 7: M2-9 SurA Fraction 3, Lane 8: Periplasmic fluid from WT SurA, Lane 9: WT SurA Wash, Lane 10: WT SurA Fraction 3

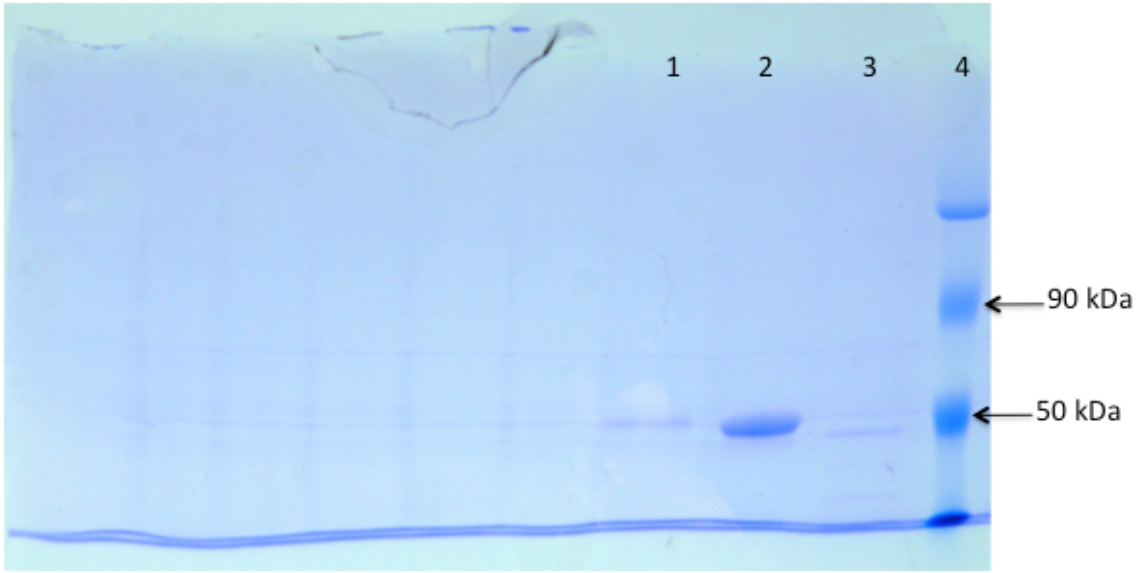

Figure 3.13 SDS PAGE gel of purified M2-7 SurA. From left: Lane 1: M2-7 Fraction 3, Lane 2: M2-7 Fraction 1, Lane 3: M2-7SurA Wash, Lane 4: Molecular Weight Marker

Following expression and purification, all three SurA mutant proteins were analyzed by secondary CD, tertiary CD, and fluorescence spectroscopy for relative structural changes. From the secondary CD plot, there are obvious differences between the mutant proteins and WT SurA. Surprisingly, there were large differences between the spectra. All mutants still maintain moderate levels of alpha-helix content, but certainly some was lost with the substitutions made at the C-terminus. The M2-9 is most different from the WT, with helical content dropped nearly 50%. And yet, the mutant is completely functional as compared to the WT SurA. Interestingly, we found that M2-9 formed a dimer band in SDS-PAGE, linked together via a Cys in the C-terminus. We speculate that SurA might need to function as a dimer, and the presence of Cys in M2-9 stabilized the dimer, and therefore improved the activity of the mutant.

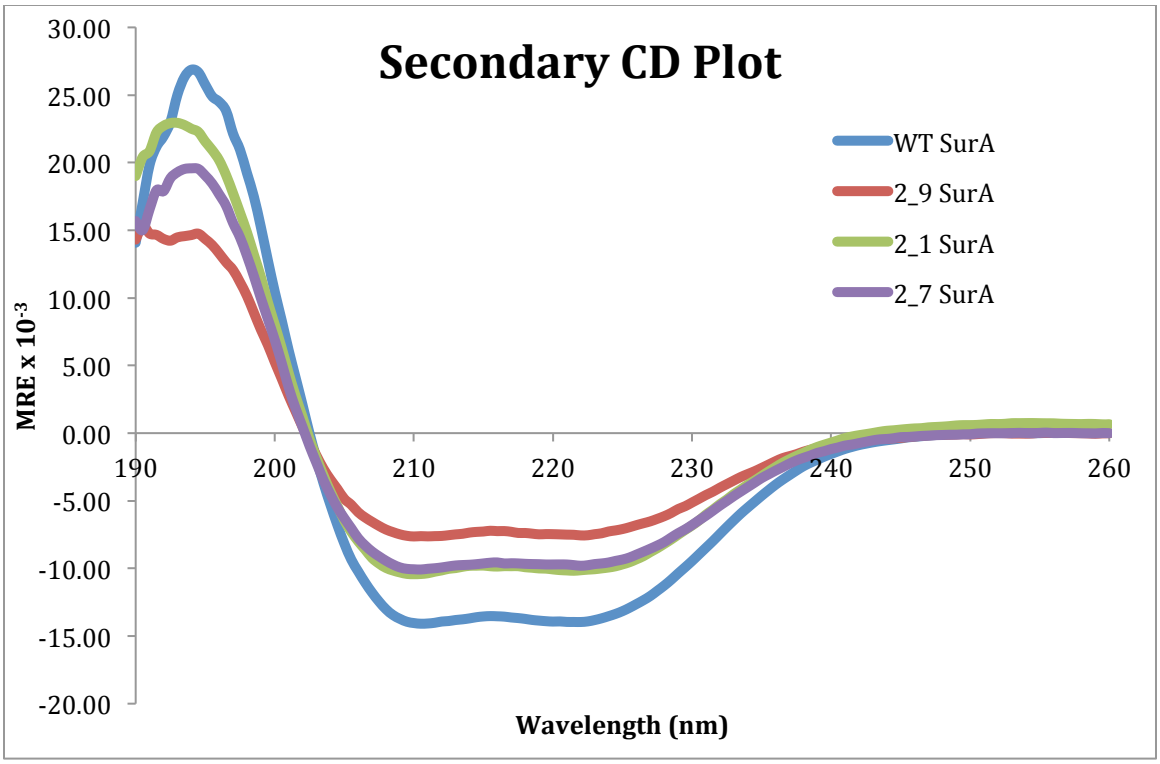

Figure 3.14 Secondary CD plot of WT with mutants $SurA_{M2\text{-}1}$, $SurA_{M2\text{-}7}$, and $SurA_{M2\text{-}9}$. All traces were normalized with to concentration of each protein. MRE = Mean Residue Ellipticity (deg*cm^2/dmol)

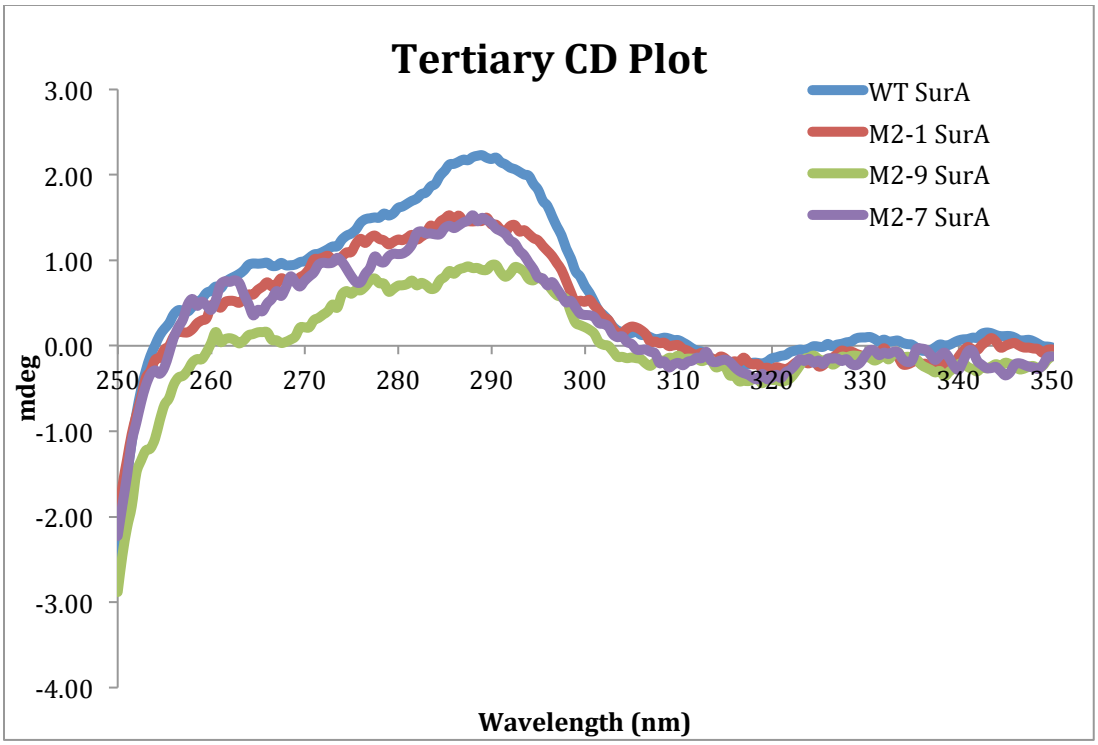

Figure 3.15 Tertiary CD plot of WT SurA with mutants $SurA_{M2-1}$, $SurA_{M2-7}$, and $SurA_{M2-9}$. All traces were normalized to the concentration of each protein.

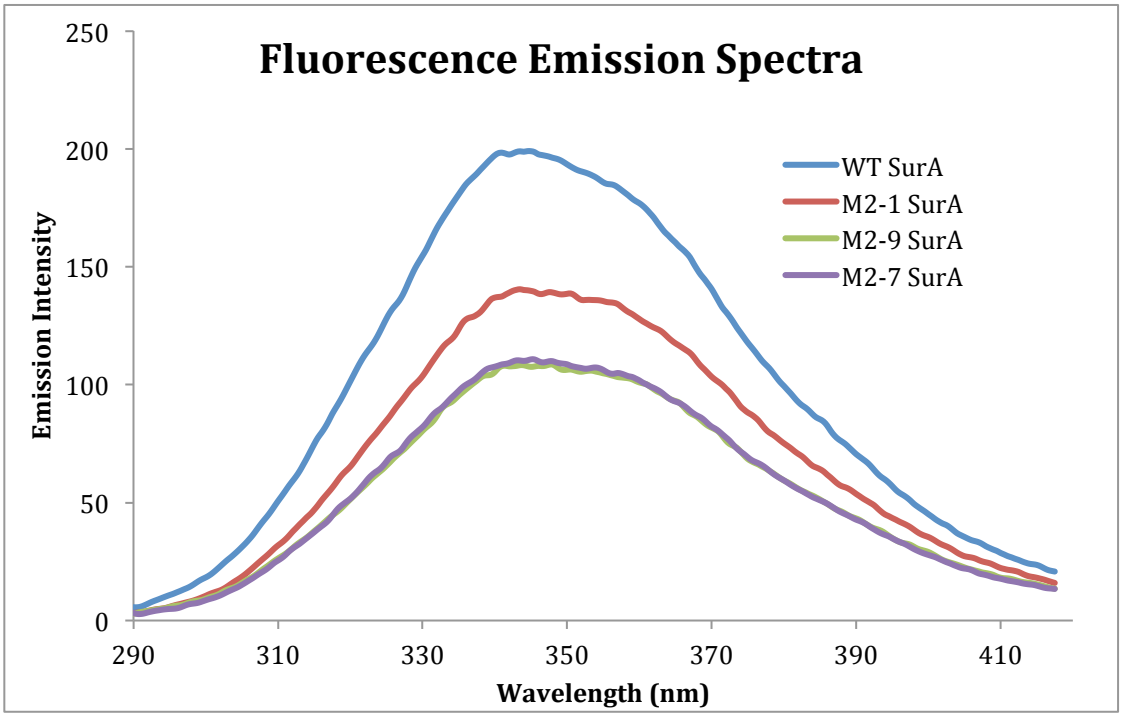

Figure 3.16 Fluorescence emission spectra of WT SurA with mutants $SurA_{M2-1,}$ $SurA_{M2-7}$, and $SurA_{M2-9}$. All traces were normalized to the concentration of each protein. Excitation wavelength was 280 nm.

The tertiary CD plots and fluorescence emission curves reveal more clearly the differences between WT SurA and the SurA mutants. Just as with the secondary CD analysis, $SurA_{M2-9}$ was the most different from WT in both the tertiary CD and fluorescence emission spectra. The peak shift and difference in intensity in the tertiary CD plots of the proteins indicates a change of three-dimensional structure. Tertiary CD can indicate relative changes in three-dimensional structure between samples, but is not precise in determining what those changes are. Thus, it can only be concluded that the mutants are, to some degree, structurally dissimilar to the WT SurA. The maximum emission peaks for WT SurA, $SurA_{M2-1}$, $SurA_{M2-7}$, $SurA_{M2-9}$ are 344.5, 343.5, 345, and 348, respectively. $SurA_{M2-9}$ is red-shifted 3.5 nm with respect to WT SurA, while the other two are shifted by only 1 nm, at the most. From the putative model for intrinsic tryptophan fluorescence emission, it can be postulated that one or more of the tryptophan residues in $SurA_{M2-9}$ is more exposed than WT SurA. Although these changes are slight, $SurA_{M2-9}$ remains the most structurally distant species from WT SurA. It is somewhat surprising that $SurA_{M2-9}$ and $SurA_{M2-1}$ show so much difference from WT, because both proteins were more active than WT SurA in the novobiocin MIC assay, and $SurA_{M2-9}$ was comparable in the SDS-EDTA MIC assay. A few possibilities could explain this: 1) The techniques used to characterize the protein structure are sensitive to changes which are not important

for SurA function, 2) SurA can accommodate a wide range of slight structural changes, which retain or even enhance function, or 3) The changes observed are artifacts and not indicative of the actual structure of the proteins. The second possibility seems most plausible due to number of active SurA mutants that were generated and the percentage that recover function from the truncated SurA template. Since there is virtually no pattern in the primary sequence of the last 10 residues of the SurA mutants, one might conclude that it is unlikely that there exists a pattern in the resulting secondary structures. However, this is not easy to predict. The last 10 residues of WT SurA form a beta-strand which associates with an N-terminal beta strand to form anti-parallel beta sheet. The propensity for primary sequences to form beta-strands is believed to be modulated primarily by the local environment, and the side chains are thought to be important for stabilizing the beta-strand [Baldi and Chang 2005]. Thus, an alternate hypothesis is that there are many different sequences that can form the beta-strand at the C-terminus based on the local environment, but some are stabilized to a higher degree than others. A decrease in the stability of the beta sheet between the N and C-termini could alter the tertiary structure of the protein, thereby accounting for the differences seen in figures above. From the C-terminal truncation results (Table 3.1), it was determined that the last 10 residues are vital for function, thus the presence of an interacting beta-strand is necessary at the C-terminus. However, we postulate that a slight decrease in the stability of this interaction is not deleterious to the function of the protein.

To further characterize the stability of the proteins *in vitro*, each of the SurA mutants were subjected to limited trypsin digestion. Trypsin is a protease that cleaves polypeptides at the carboxyl side of arginine and lysine residues. All protein levels were normalized and trypsin was allowed to digest for 61 minutes before heat inactivation along with SDS and DTT (a reducing agent for disulfide bonds). Samples were removed at 1 minute and every 10 minutes after that until 61 minutes had passed (Figures 3.17-3.20). Figures 3.22 and 3.23 show the same digestion, except samples were removed before and immediately following trypsin addition.

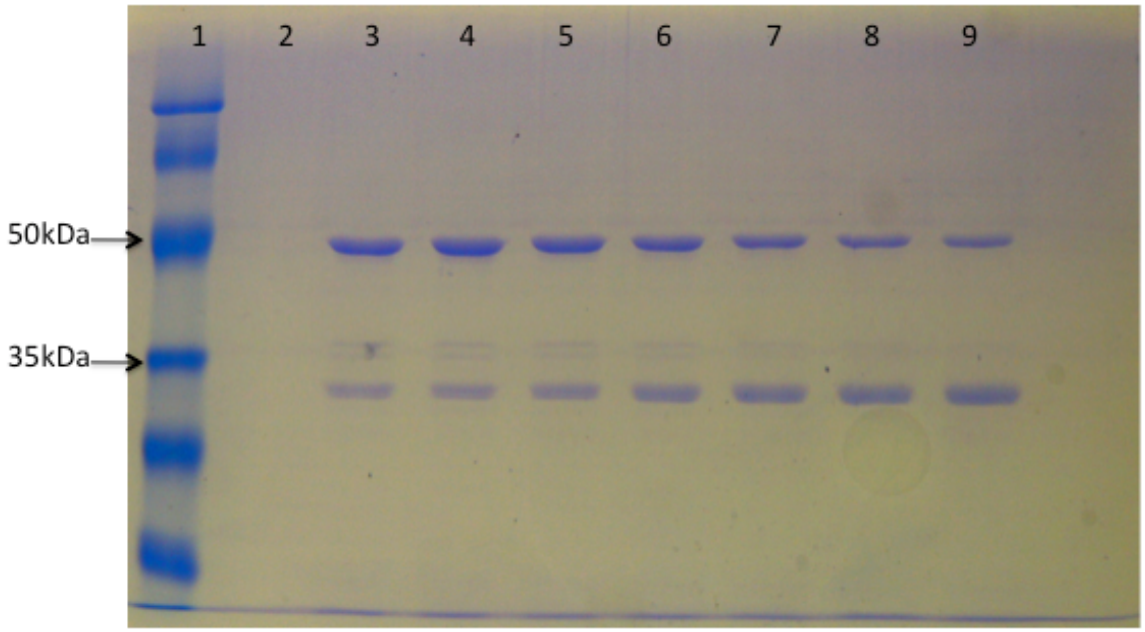

Figure 3.17 Trypsin digestion of WT SurA on 15% SDS gel. From left, Lane 1: Molecular Weight marker. Lane 2: Trypsin without SurA. Lanes 3-9: SurA digestion starting at 1 minute and increasing at 10 minute intervals to 61 minutes. Lane 10: Trypsin without SurA.

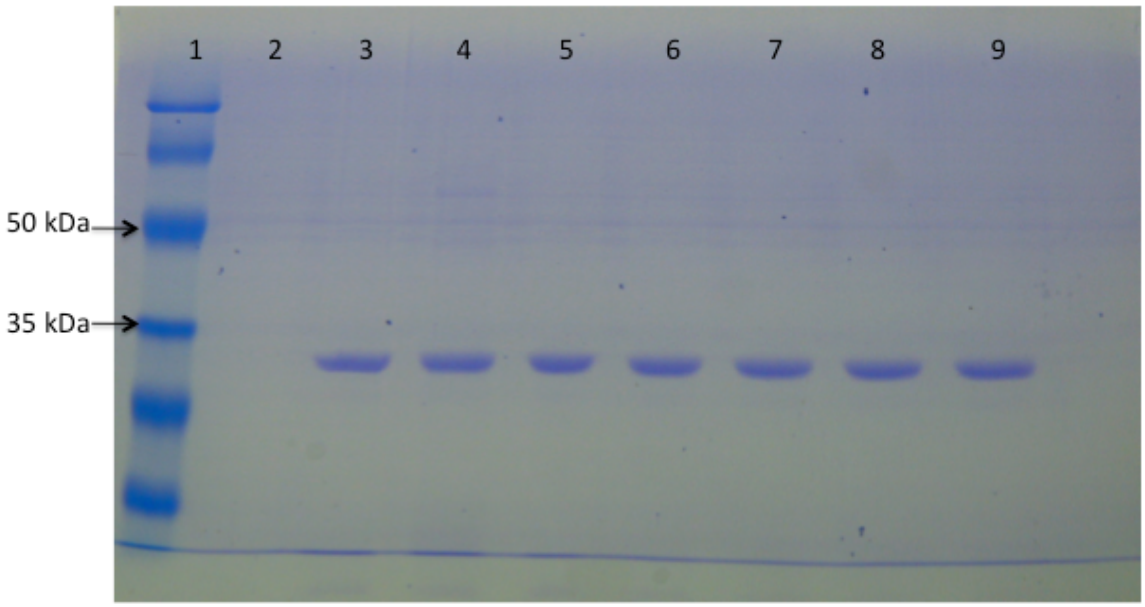

Figure 3.18 Trypsin digestion of M2-1 SurA on 15% SDS gel. Lane 1: Molecular Weight marker. Lane 2: Trypsin without SurA. Lanes 3-9: SurA digestion starting at 1 minute and increasing at 10 minute intervals to 61 minutes.

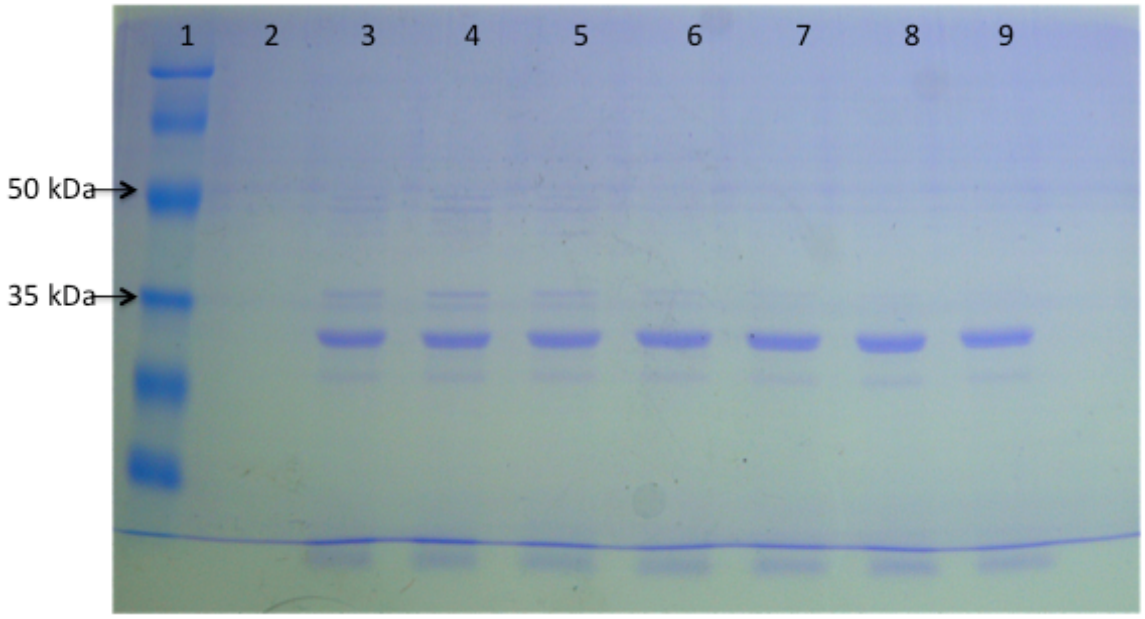

Figure 3.19 Trypsin digestion of M2-9 SurA on 15% SDS gel. Lane 1: Molecular Weight marker. Lane 2: Trypsin without SurA. Lanes 3-9: SurA digestion starting at 1 minute and increasing at 10 minute intervals to 61 minutes.

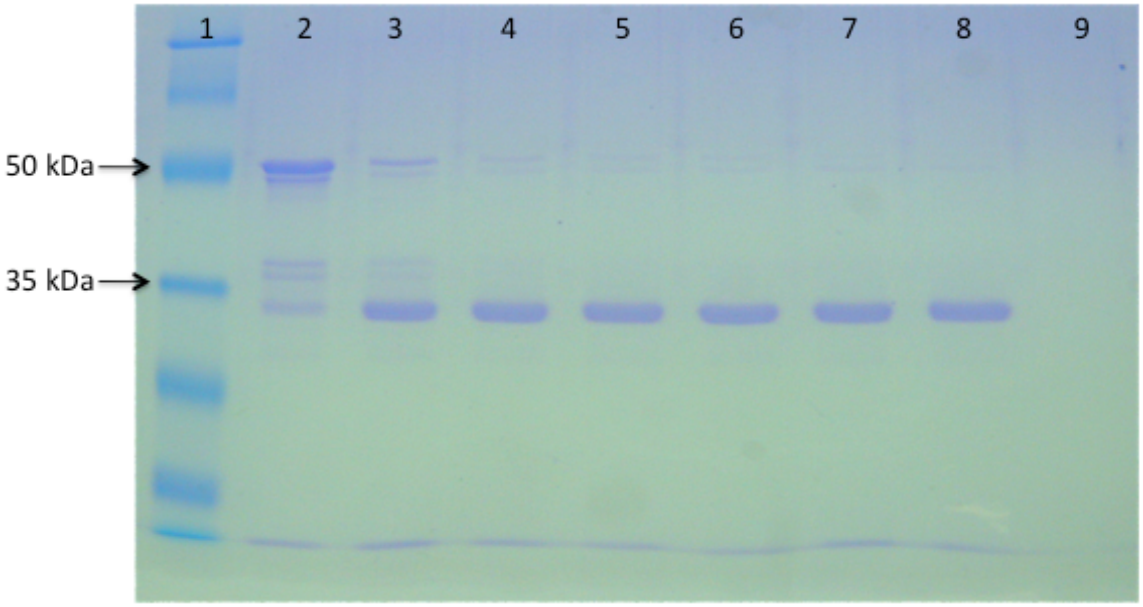

Figure 3.20 Trypsin digestion of M2-7 SurA on 15% SDS gel. Lane 1: Molecular Weight marker. Lanes 2-8: SurA digestion starting at 1 minute and increasing at 10 minute intervals to 61 minutes. Lane 9: Trypsin without SurA.

Limited trypsin digestion of WT SurA and SurA mutants revealed surprising results. The two most active mutants ($SurA_{M2-1}$ and $SurA_{M2-9}$) are extremely susceptible to cleavage by trypsin, with almost complete digestion occurring within 1 minute (Fig. 3.18 and 3.19). $SurA_{M2-7}$ lasted slightly longer, with complete digestion occurring between 10 and 20 minutes (Fig. 3.20). WT SurA was susceptible to cleavage, but even after 61 minutes, a significant band was still visible, indicating incomplete digestion. To rule out the possibility that the protein was degraded prior to digestion, a separate experiment was performed where the samples were removed just prior to trypsin addition (Fig. 3.22 and 3.23).

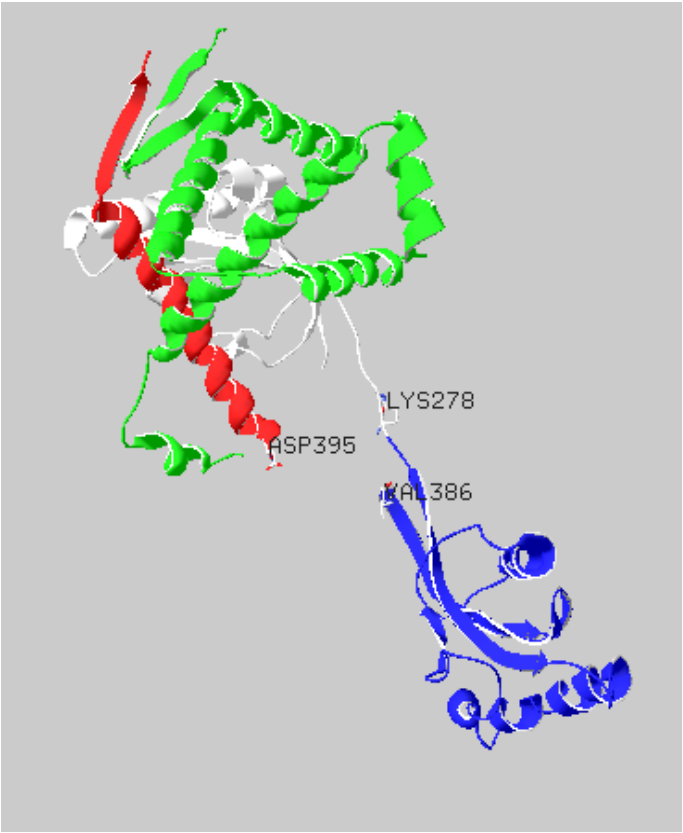

Figure 3.21 Crystal structure of full-length SurA with proposed cutting sites, K(278) and K(388), for trypsin. The region between V(386) and D(395) was not resolved in the crystal structure, but it is presumed that K(388) lies very close to K(278) on the adjacent linker polypeptide. Structure was generated from the Swiss Protein Data Bank, ID: 1M5Y [Bitto, E. and D. B. McKay (2002)].

Perhaps the most interesting finding of this experiment is the fact that all of the SurA mutants were digested to the same sized fragment, which was resistant to further digestion. Consequently, there must be an exposed fragment of ~ 16 kDa

connected to a tightly packed fragment that is ~ 30 kDa. We propose that the cutting site is K(278) , which is positioned in the middle of one of the linker regions between the core domain and the satellite P2 domain(Fig. 3.21). If spliced, the resulting molecular weights of the core domain and P2 domain would be 29.1 kDa and 16.8 kDa, respectively. The primary fragment on the gel runs in the middle of the 35 kDa marker and 25kDa marker, and the smaller fragment can be seen most easily on Figure 19, just below the dye band. Although these fragments are most likely the core domain and satellite P2 domain, only analysis by mass spectrometry can confirm this with accuracy. There are a few other fragments present on the gels, particularly in the case of Figure 19. This is likely due to the presence of dimers. DTT was not added until after the digestion was completed, so these dimers were present during the digestion.

This finding, although very interesting, is difficult to explain. When substitution at the C-terminus occurs, in all three cases, SurA becomes very susceptible to trypsin digestion. Taken alone, this would suggest that the protein is very unstable *in vivo*, and in fact, it might be. Nevertheless, if our hypothesis regarding the cutting site is correct, then it would not matter if SurA is digested easily *in vivo*, because SurA(ΔP2) is just as active as full-length SurA [Hennecke, G., J. Nolte, et al. (2005), Behrens, S., R. Maier, et al. (2001)]. This finding corroborates the results from tertiary CD and fluorescence spectroscopy by showing that the mutant proteins are in fact, structurally different. The substitutions at the C-terminus could loosen the interaction between the N and C-termini, creating a more

flexible structure and leaving the P2 domain tethered, but very accessible to digestion.

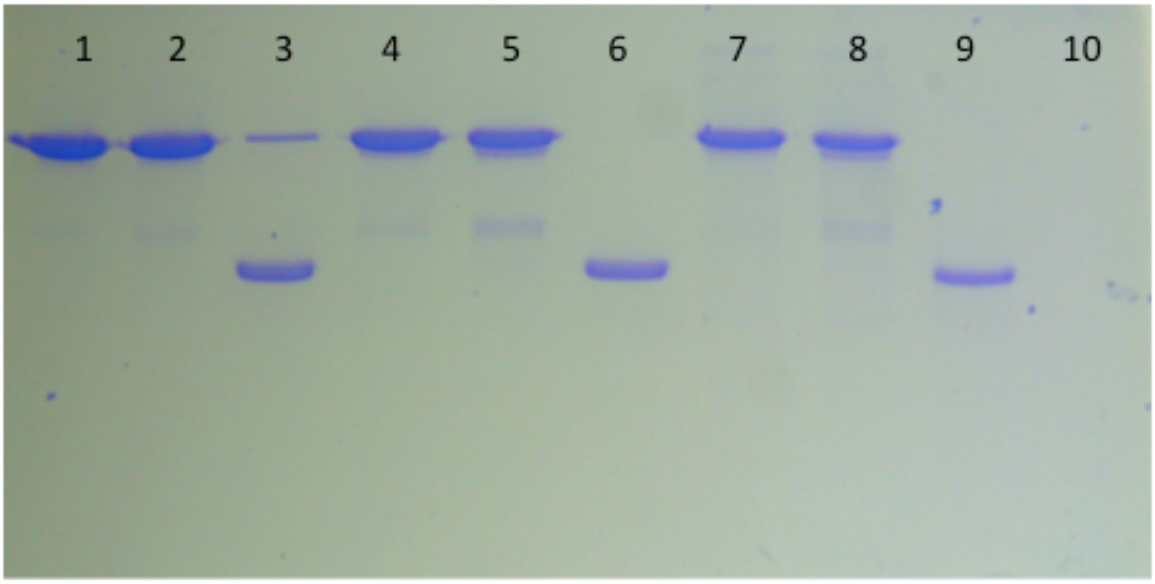

Figure 3.22 Trypsin digestion of WT, M2-1, and M2-9 SurA proteins for 60 minutes on 20% SDS gel. From left, Lane 1: WT SurA, no trypsin, Lane 2: WT SurA at 30 s, Lane 3: WT SurA at 60 min., Lane 4: M2-1 SurA, no trypsin, Lane 5: M2-1 SurA at 30 s, Lane 6: M2-1 SurA at 60 min., Lane 7: M2-9 SurA, no trypsin, Lane 8: M2-9 SurA at 30 s, Lane 9: M2-9 SurA at 60 min., Lane 10: Trypsin without SurA.

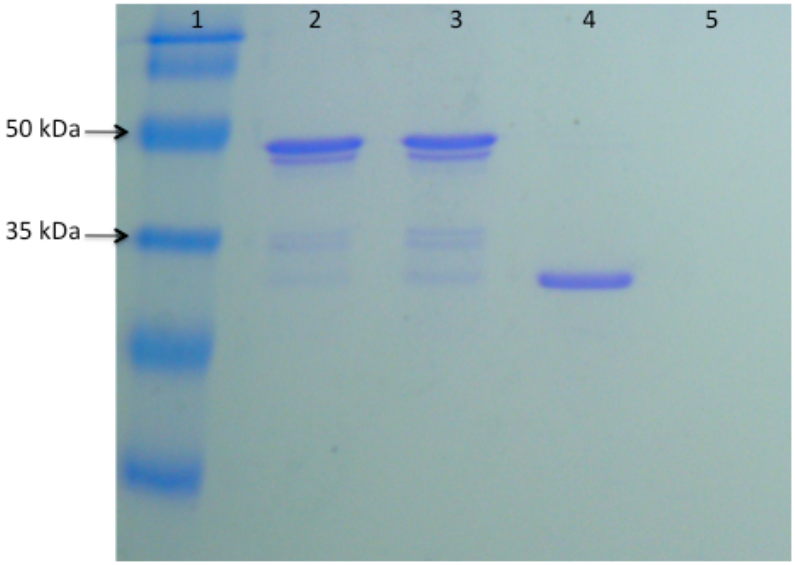

Figure 3.23 Trypsin digestion of M2-7 SurA for 60 minutes on 20% SDS gel. From left, Lane 1: Molecular Weight marker, Lane 2: M2-7 SurA, no trypsin, Lane 3: M2-7 SurA at 30 s, Lane 4: M2-7 SurA at 60 min., Lane 5: Trypsin without SurA.

Chapter IV: Conclusions

The purpose of this thesis work was to identify key residues at the C-terminus of SurA involved in the chaperone function. The first half of the study focused on site-directed mutagenesis of conserved residues in SurA. After many mutations that were ineffective at suppressing the function of SurA, one mutation was found, E(408):A, that caused almost a 3-fold reduction in MIC of novobiocin compared to wild-type. Characterization of $SurA_{E408A}$ by CD, fluorescence emission, and fluorescence polarization revealed only slight changes in the structure and function when compared with WT SurA. We postulate that the difference in MIC could be explained by differences in the stability of the two proteins. Since an in-depth study on the stability of these proteins was not conducted, we can only hypothesize that this is the case.

A small, intermediate study was conducted that centered around the role of aromatic residues in SurA. We generated 16 mutants of single, double, and triple mutations at every aromatic residue on SurA. Only two (Y125,398,422:A and Y128,398,422:A) were found that significantly inhibited SurA function. We propose that this is due to destabilization of the protein when the single mutations are combined.

The second half of this work focused on substituting the last 10 residues of SurA with a random peptide library of the same length. The last 10 residues are vital for SurA function and, when truncated, result in a non-functional protein. We found that ~1.5% of all possible primary amino acid sequences will recover SurA function

to at least 50% of wild-type activity. Twenty-six mutants were selected at low concentrations of novobiocin, sequenced, and tested for MIC with both novobiocin and SDS-EDTA. We found that there is no pattern in the sequences of mutants that recovered SurA function. The only position that showed any sequence homology among the mutants was the valine at position 5 in the 10-mer. The most active mutants all have either valine, or close substitutes, methionine or isoleucine at that position. However, it is unlikely that a single mutation at this position would be sufficient to generate an inactive SurA protein, because other mutants, which are only slightly less active, have other substitutions at that position. In conclusion, we feel that SurA is a very flexible protein, capable of accommodating many changes to its sequence, while still maintaining activity. Other proteins, like SurA, that have a wide range of substrates, might also behave in a similar way when changes are introduced to the primary sequence.

References

Baron, S. (1996), Medical Microbiology, 4 edition

Behrens, S.; Maier, R.; de Cock, H.; Schmid, F. X.; Gross, C. A., The SurA periplasmic PPlase lacking its parvulin domains functions in vivo and has chaperone activity. *Embo Journal* **2001,** *20* (1-2), 285-294.

Bernstein, H.D. (2000) The biogenesis and assembly of bacterial membrane proteins. *Curr. Opin. Microbiol.* 3, 203–209.

Bitto, E.; McKay, D. B., Binding of phage-display-selected peptides to the periplasmic chaperone protein SurA mimics binding of unfolded outer membrane proteins. *Febs Letters* **2004,** *568* (1-3), 94-98

Bitto, E.; McKay, D. B., Crystallographic structure of SurA, a molecular chaperone that facilitates folding of outer membrane porins. *Structure* **2002,** *10* (11), 1489-1498.

Bitto, E.; McKay, D. B., The periplasmic molecular chaperone protein SurA binds a peptide motif that is characteristic of integral outer membrane proteins. *J. Biol. Chem.* **2003,** *278* (49), 49316-49322.

Bos, M. P.; Robert, V.; Tommassen, J., Biogenesis of the gram-negative bacterial outer membrane. *Annual Review of Microbiology* **2007,** *61*, 191-214.

Boxer, D. H.; Zhang, H.; Gourley, D. G.; Hunter, W. N.; Kelly, S. M.; Price, N. C., Sensing of remote oxyanion binding at the DNA binding domain of the molybdate-dependent transcriptional regulator, ModE. *Org. Biomol. Chem.* **2004,** *2* (19), 2829-2837.

Bulieris P. V., Behrens S., Holst O. and Kleinschmidt J. H. (2003) Folding and insertion of the outer membrane protein OmpA is assisted by the chaperone Skp and by lipopolysaccharide. *J. Biol. Chem.* 278, 9092–9099.

Chen R. and Henning U. (1996) A periplasmic protein (Skp) of *Escherichia coli* selectively binds a class of outer membrane proteins. *Mol. Microbiol.* 19, 1287–1294.

Cheng, J. L.; Baldi, P., Three-stage prediction of protein beta-sheets by neural networks, alignments and graph algorithms. *Bioinformatics* **2005,** *21*, I75-I84.

Denoncin, K.; Vertommen, D.; Paek, E.; Collet, J. F., The Protein-disulfide Isomerase DsbC Cooperates with SurA and DsbA in the Assembly of the Essential beta-Barrel Protein LptD. *J. Biol. Chem.* **2010** *285* (38), 29425-29433.

Duguay A. R., Silhavy T. J. (2004) Quality control in the bacterial periplasm. *Biochim. Biophys. Acta* 1694, 121– 134.

Fischbach, M. A.; Walsh, C. T., Antibiotics for Emerging Pathogens. *Science* **2009,** *325* (5944), 1089-1093.

Gold, V. A. M., Duong, F. and Collinson, I. (2007). "Structure and Function of the Bacterial Sec Translocon." *Mol. Membr. Biol.* 24, 387-394.

Harms N., Koningstein G., Dontje W., Muller M., Oudega B., Luirink J. and de Cock H. (2001) The early interaction of the outer membrane protein phoe with the periplasmic chaperone Skp occurs at the cytoplasmic membrane. *J. Biol. Chem.* 276, 18804–18811.

Hennecke, G.; Nolte, J.; Volkmer-Engert, R.; Schneider-Mergener, J.; Behrens, S., The periplasmic chaperone SurA exploits two features characteristic of integral outer membrane proteins for selective substrate recognition. *J. Biol. Chem.* **2005,** *280* (25), 23540-23548.

Jameson, D. M.; Seifried, S. E., Quantification of protein-protein interactions using fluorescence polarization. *Methods-a Companion to Methods in Enzymology* **1999,** *19* (2), 222-233.

Jaud, S., Fernandez-Vidal, M., Nilsson, I., Meindi-Beinker, N. M., Hubner, N. C., Tobias, D. J., von Heijne, G. and White, S. H. (2009). "Insertion of Short Transme3mbrane Helices by the Sec61 Translocon." *Proc. Natl. Acad. Sci. U. S. A.* 106, 11588-11593.

Johnson, W. C., PROTEIN SECONDARY STRUCTURE AND CIRCULAR-DICHROISM - A PRACTICAL GUIDE. *Proteins-Structure Function and Genetics* **1990,** *7* (3), 205-214.

Justice, S. S.; Hunstad, D. A.; Harper, J. R.; Duguay, A. R.; Pinkner, J. S.; Bann, J.; Frieden, C.; Silhavy, T. J.; Hultgren, S. J., Periplasmic peptidyl prolyl cis-trans isomerases are not essential for viability, but SurA is required for pilus biogenesis in Escherichia coli. *Journal of Bacteriology* **2005,** *187* (22), 7680-7686.

Kelly, S. M.; Jess, T. J.; Price, N. C., How to study proteins by circular dichroism. *Biochimica Et Biophysica Acta-Proteins and Proteomics* **2005,** *1751* (2), 119-139.

Lazar, S. W.; Kolter, R., SurA assists the folding of Escherichia coli outer membrane proteins. *Journal of Bacteriology* **1996,** *178* (6), 1770-1773.

Lizak, B., Csala, M., Benedetti, A. and Banhegyi, G. (2008). "The Translocon and the Non-Specific Transport of Small Molecules in the Endoplasmic Reticulum (Review)." *Mol. Membr. Biol.* 25, 95-101.

Missiakas, D., Betton, J. M. and Raina, S. (1996) New components of protein folding in extracytoplasmic compartments of *Escherichia coli* SurA, FkpA and Skp/OmpH. *Mol. Microbiol.* 21, 871–884.

Mitscher, L. A.; Pillai, S. P.; Gentry, E. J.; Shankel, D. M., Multiple drug resistance. *Medicinal Research Reviews* **1999,** *19* (6), 477-496.

Mogensen J. E., and Otzen D. E. (2005) Interactions between folding factors and bacterial outer membrane proteins. *Mol. Microbiol.* 57, 326–346.

Palomino, C.; Marin, E.; Angel Fernandez, L., The Fimbrial Usher FimD Follows the SurA-BamB Pathway for Its Assembly in the Outer Membrane of Escherichia coli. *Journal of Bacteriology* **2011** *193* (19), 5222-5230.

Pluckthun, A, Wulfing, C. Protein folding in the periplasm of *Escherichia coli*. Moiecuiar Microbioiogy (1994) 12(5), 685-692.

Qu, J.; Mayer, C.; Behrens, S.; Holst, O.; Kleinschmidt, J. H., The trimeric periplasmic chaperone skp of Escherichia coli forms 1 : 1 complexes with outer membrane proteins via hydrophobic and electrostatic interactions. *J. Mol. Biol.* **2007,** *374* (1), 91-105.

Rizzitello, A. E., J. R. Harper, et al. (2001). "Genetic evidence for parallel pathways of chaperone activity in the periplasm of Escherichia coli." Journal of Bacteriology **183**(23): 6794-6800.

Rouviere, P. E.; Gross, C. A., SurA, a periplasmic protein with peptidyl-prolyl isomerase activity, participates in the assembly of outer membrane porins. *Genes & Development* **1996,** *10* (24), 3170-3182.

Ruiz N., Kahne D., Silhavy T.J. (2006) Advances in understanding bacterial outer-membrane biogenesis *Nat. Rev. Microbiol.* 4, 57–66.

Rusch, S. L. and Kendall, D. A. (2007). "Oligomeric States of the Seca and SecyegCore Components of the Bacterial Sec Translocon." *Biochim. Biophys. Acta Biomembranes* 1768, 5-12.

Schafer U., Beck K. and Muller M. (1999) Skp, a molecular chaperone of gram-negative bacteria, is required for the formation of soluble periplasmic intermediates of outer membrane proteins. *J. Biol. Chem.* 274, 24567–24574.

Sijbrandi, R., Urbanus, M.L., ten Hagen-Jongman, C.M., Bernstein, H.D., Oudega, B., Otto, B.R., and Luirink, J. (2003) Signal recognition particle (SRP)-mediated

targeting and Sec-dependent translocation of an extracellular *Escherichia coli* protein. *J. Biol. Chem.* 278, 4654–4659.

Sklar J. G., Wu T., Kahne D., and Silhavy T. J. (2007) Defining the roles of the periplasmic chaperones SurA, Skp, and DegP in *Escherichia coli*. *Genes Dev.* 21, 2473-2484.

Stymest K. H., and Klappa P. (2008) The periplasmic peptidyl prolyl cis-trans isomerases PpiD and SurA have partially overlapping substrate specificities. *FEBS J.* 275, 3470-3479.

Teale, F.J.W. The ultraviolet fluorescence of proteins in neutral solution. *Journal of Biochemistry* **1960,** 76, 381-388.

Tormo, A.; Almiron, M.; Kolter, R., SURA, AN ESCHERICHIA-COLI GENE ESSENTIAL FOR SURVIVAL IN STATIONARY PHASE. *Journal of Bacteriology* **1990,** *172* (8), 4339-4347

Velayati, A. A.; Masjedi, M. R.; Farnia, P.; Tabarsi, P.; Ghanavi, J.; ZiaZarifi, A. H.; Hoffner, S. E., Emergence of New Forms of Totally Drug-Resistant Tuberculosis Bacilli Super Extensively Drug-Resistant Tuberculosis or Totally Drug-Resistant Strains in Iran. *Chest* **2009,** *136* (2), 420-425

Vivian, J. T.; Callis, P. R., Mechanisms of tryptophan fluorescence shifts in proteins. *Biophys. J.* **2001,** *80* (5), 2093-2109.

Watts, K. M.; Hunstad, D. A., Components of SurA Required for Outer Membrane Biogenesis in Uropathogenic Escherichia coli. *PLoS One* **2008,** *3* (10), 6.

Woody, A. Y. M.; Woody, R. W., Individual tyrosine side-chain contributions to circular dichroism of ribonuclease. *Biopolymers* **2003,** *72* (6), 500-513.

Xu, X. H.; Wang, S. Y.; Hu, Y. X.; McKay, D. B.The Periplasmic Bacterial Molecular Chaperone SurA Adapts its Structure to Bind Peptides in Different Conformations to Assert a Sequence Preference for Aromatic Residues. (2007) J.Mol.Biol. **373:** 367-381.

Vita

Brent Ferrell

Personal

Date of Birth: September 9th, 1985

Place of Birth: Raleigh, NC

Academic Degrees

M.S. Chemistry, University of Kentucky, Lexington, KY	2009-Present
B.S. Biochemistry, Lee University, Cleveland, TN	2005-2008

Professional Experience

Teaching Assistant, University of Kentucky	2010-2011
Adjunct Instructor, Georgetown College	2012

Honors/Awards

Recipient, NASA EPSCoR Graduate Fellowship	2009-2010
Recipient, NSF-REU Undergraduate Research Fellowship	2009
Graduate, Cum Laude, Lee University	2008